Deduplication on Scrambled Enormous Information in Cloud

Dr.R.Balakrishna, Prasad.A.Y, Monika R, Megha S

Dr.R.Balakrishna,Principal,Rajarajeshwari college of Engineering,Bangalore.

Prasad.A.Y, Asst Professor,Rajarajeshwari college of Engineering,Bangalore.

Monika R, VIII Sem,ISE,Dept, Rajarajeshwari college of Engineering,Bangalore.

Megha S, VIII Sem,ISE, Dept.Rajarajeshwari college of Engineering,Bangalore.

Abstract—**Cloud figuring offers another method for administration arrangement by modifying different assets over the Internet. Themost imperative and prominent cloud administration is information stockpiling. With a specific end goal to protect the security of information holders, information are frequently put away in cloud in a scrambled shape. In any case, scrambled information present new difficulties for cloud information deduplication, which winds up noticeably pivotal for huge information stockpiling and preparing in cloud. Conventional deduplication plans can't take a shot at scrambled data. Existing arrangements of encoded information deduplication experience the ill effects of security shortcoming. They can't adaptably bolster information get to control and renouncement. Subsequently, few of them can be promptly sent by and by. In this paper, we propose a plan to deduplicate encoded information put away in cloud in view of proprietorship test and intermediary re-encryption.It coordinates cloud information deduplication with get to control. We assess its execution in light of broad examination and PC recreations. The outcomes demonstrate the unrivaled proficiency and viability of the plan for potential functional organization, particularly for huge information deduplication in distributed storage.**

Index Terms—**Access Control, Big Data, Cloud Computing, Data Deduplication, Proxy Re-encryption.**

INTRODUCTION

Boisterous processing offers another method for Information Technology benefits by adjusting different assets (e.g., capacity, registering) and giving them to clients in light of their requests. Distributed computing gives a major asset pool by connecting system assets together. It has alluring properties, for example, versatility, flexibility, adaptation to non-critical failure, and pay-per-utilize. Along these lines, it has turned into a promising administration platform. The most critical and famous cloud administration is information capacity benefit. Cloud clients transfer individual or private information to the server farm of a Cloud Service Provider (CSP) and enable it to keep up these information. Since interruptions and assaults towards delicate information at CSP are not maintain a strategic distance from capable, it is reasonable to accept that CSP can't be completely trusted by cloud clients. Additionally, the loss of control over their very own information [44, 45] prompts high information security dangers, particularly information protection spillages. Because of the quick advancement of information mining and different examination advances, the security issue ends up noticeably genuine. Henceforth, a decent practice is to just outsource scrambled information to the cloud keeping in mind the end goal to guarantee information security and client protection. In any case, the same or distinctive clients may transfer copied information in scrambled shape to CSP, particularly for situations where incredibly squanders organize assets, expends a great deal of vitality, and muddles information administration. The improvement of various administrations additionally makes it dire to convey productive asset administration system. Thusly, deduplication winds up noticeably basic for huge information stockpiling and preparing in the cloud. Deduplication has demonstrated to accomplish highcost funds, e.g., decreasing up to 90-95% capacity requirements for reinforcement applications and up to 68% in standard record frameworks. Clearly, the reserve funds, which can be passed back specifically or in a roundabout way to cloud clients, are critical to the financial matters of cloud business. Instructions to oversee encoded information capacity with deduplication in a productive way is a useful issue. Be that as it may, current mechanical deduplication arrangements can't deal with scrambled information. Existing answers for deduplication experience the ill effects of bruteforce assaults [7, 11]. They can't adaptably bolster information get to control and renouncement at the same time[16, 18-20]. Most existing arrangements can't guarantee dependability, security and protection with sound execution [23].In practice, it is difficult to enable information holders to oversee deduplication because of various reasons. To start with, information holders may not be constantly on the web or accessible for such an administration, which could bring about capacity delay. Second, deduplication could turn out to be excessively confused as far as interchanges and calculations to include information holders into deduplication prepare. Third, it might encroach the protection of information holders during the time spent finding copied information. Forward, an information holder may have no clue how to issue information get to rights or deduplication keys to a client in a few circumstances when it does not know other information holders because of information super dispersion. Consequently, CSP can't coordinate with information holders on information stockpiling deduplication much of the time. In this paper, we propose a plan in view of information proprietorship test and Proxy Re-Encryption (PRE) to oversee encoded information stockpiling with deduplication. We expect to comprehend the issue of deduplication in the circumstance where the information holder is not accessible or hard to get included. Then, the execution of information deduplication in our plan is not impacted by the span of information, therefore pertinent for huge information. In particular, the commitments of this paper can be condensed as beneath:

- We rouse to spare distributed storage and safeguard the protection of information holders by proposing a plan to oversee scrambled information stockpiling with deduplication. Our plan can adaptably bolster information sharing with deduplication notwithstanding when the information holder is disconnected, and it doesn't encroach the security of information holders.
- We propose a compelling way to deal with confirm information possession and check copy stockpiling with secure test and huge information bolster.
- We incorporate cloud information deduplication with information get to control basically, in this manner accommodating information deduplication and encryption.
- We demonstrate the security and survey the execution of the proposed plot through examination and recreation. The outcomes demonstrate its proficiency, adequacy and appropriateness. Whatever remains of the paper is sorted out as takes after. Area 2 gives a short outline of related work. Segment 3 presents framework and security models, preliminaries and documentation. Segment 4 gives the nitty gritty portrayal of our plot, trailed by security investigation and execution assessment in Section 5. At long last, a conclusion is exhibited in the last segment.

2 RELATED WORK

2.1 Encrypted Data Deduplication

Distributed storage specialist co-ops, for example, Dropbox [2], Google Drive [3], Mozy [4], and others perform deduplication to spare space by just putting away one duplicate of each file transferred. Notwithstanding, if customers customarily encode their information, stockpiling investment funds by deduplication are completely lost. This is on account of the scrambled information are spared as different substance by applying distinctive encryption keys. Existing modern arrangements f trouble in scrambled information deduplication. For instance, DeDu [17] is a productive deduplication framework, yet it can't deal with encoded information. Accommodating deduplication and customer side encryption is a dynamic research point [1]. Message Bolted Encryption (MLE) int closures to tackle this issue [5]. The most prominent indication of MLE is Convergent Encryption (CE), presented by Douceur et al. [6] and others [7, 11, 12]. CE was utilized inside a wide assortment of business furthermore, look into capacity benefit frameworks. Letting M be a record's information , a customer initially figures a key $K \leftarrow H(M)$ by applying a cryptographic hash work H to M , and afterward figures ciphertext $C \leftarrow E(K,M)$ through a deterministic symmetric encryption plot. A moment customer B scrambling the same document M will create the same C , empowering deduplication. Be that as it may, CE is liable to a natural security confinement, to be specific , powerlessness to disconnected beast drive word reference assaults [13, 14]. Realizing that the objective information M underlying the target ciphertext C is drawn from a word reference $S=\{M1,...,Mn\}$ of size n , an aggressor can recuperate M in the time for $n=|S|$ off - line encryptions: for each $i=1,...,n$, it essentially CE - scrambles Mi to get a ciphertext signified as Ci also, returns Mi to such an extent that $C=Ci$. This works because CE is deterministic and key

less. The security of CE is just possible when the objective information is drawn from a space too substantial to deplete. Another issue of CE is that it is not adaptable to support information get to control by information holders, particularly for data repudiation procedure , since it is unimaginable for information holders to produce the same new key for information re-encryption [18, 19]. A picture deduplication conspire receives two servers to accomplish irrefutability of deduplication [18]. The CE - based plan depicted in [19] combines record substance and client benefit to acquire a document token with token unforgeability. In any case, both plans directly encode information with a CE key, subsequently experience the ill-effects of the aceblem as depicted previously. To oppose the assault of manipulation of information identifier, Meye et al. proposed to embrace two servers for intra - client deduplication what's more, entomb - deduplication [20]. The ciphertext C of CE is further encrypted with a client key and exchanged to the servers. Be that as it may, it doesn't manage information sharing after deduplication among various clients. ClouDedup [16] too intends to adapt to the inborn security exposures of CE, be that as it may, it can't tackle the issue created by information erasure. A information holder that expels the information from the cloud can even now get to similar information since it still knows the information encryption key if the information is not totally expelled from the cloud. Bellareet al. [1] proposed DupLESS that gives secure deduplicated capacity to oppose savage - compel assaults. In DupLESS, a gathering of associated customers (e.g., organization employees) scramble their information with the guide of a Key Server (KS) that is separate from a Storage Service (SS). Customers confirm themselves to the KS, yet don't release any data about their information to it. For whatever length of time that the KS remains out of reach to attackers, high security can be ensured. Clearly, DupLESS can't control information access of other information clients adaptably. On the other hand, a strategy - based deduplication intermediary conspire [15] was proposed however it didn't consider copied information administration (e.g., deletion what's more, proprietor administration) and did not assess conspire execution. As expressed in [21], unwavering quality, security and protection ought to be taken into contemplations when outlining a deduplication framework. The strict inertness prerequisites of essential stockpiling lead to the attention on disconnected deduplication frameworks [22]. Late reviews proposed methods to improve demonstrate reestablish execution [23]. Fu et al. [23] proposed History Mindful Rewriting (HAR) calculation to precisely recognize and modify divided lumps, which improved the reestablish execution. Kaczmarczyk et al. centered on entomb form duplication and proposed Context - Based Revamping (CBR) to enhance the reestablish execution for most recent reinforcements by moving discontinuity to more seasoned reinforcements. Another work even expert postured to relinquish deduplication to decrease the lump fracture by holder topping. In our past work, we proposed utilizing PRE for cloud information deduplication. This plan applies the hash code of information to check proprietorship with mark verification, which is shockingly unreliable in the event that H(M) is uncovered to a malignant client . In this paper, we propose another ownership confirmation approach to enhance our past work and intend to bolster enormous information deduplication in an productive way.

2.2 Information Ownership Verification and Others

Halevi et al. to start with presented the useful implementation of Proofs of Ownership (PoW) in view of Merkle tree for deduplication, which acknowledged customer side deduplication. They proposed to apply an eradication coding or, then again hash work over the first record to start with and after that utilization Merkle tree on the pre - prepared information to produce the confirmation data. While testing a prover, a verifier randomly picks a few leaves of the tree and gets the comparing kin - ways of every one of these takes off. As it were at the point when all ways are legitimate, will the verifier acknowledge the confirmation. This development can distinguish deduplication at a customer to spare arrange data transmission what's more, assurance that the customer holds a record instead of some part. Pietro et al. picked the projection of a record onto some haphazardly chose bit - positions as prooftop to understand the PoW. In any case, both plans above don't focus on information security what's more, the server for information stockpiling could be mindful of the document content . Ng et al. Adjusted the PoW to deal with the deduplication of scrambled information. This plan too creates confirmation data for deduplication check in light of Merkle trees. Each leaf esteem is created in view of a few information pieces, while each intelligent verification convention can as it were move one leaf of the Merkle tree. With a specific end goal to accomplish higher security by checking more information , the convention ought to be executed various circumstances which presents much overhead. Yang et al. too proposed a cryptographically secure what's more, efficient plot to check the responsibility for record , in which a customer demonstrates to the server that it in reality forces the whole record without transferring the document. By relying on unique spot checking, an information holder just needs to get to little yet powerful portions of the first document to create the confirmation of ownership of the first document, along these lines significantly lessening the weight of calculation on the information holder and limiting the correspondence fetched between the information holder and CSP. In the meantime, by using dynamic coefficients and haphazardly picked records of the unique documents, the plan blends the haphazardly examined segments of the first document with the dynamic coefficients to produce the one of a kind verification in each test. The work concentrates on possession evidence of the transferred information amid information deduplication. In this paper, we utilize Eillptic Curve Cryptography (ECC) to check information possession with the support of an approved gathering. Yuan and Yu endeavored to fathom the issue of supporting effective and secure information respectability auditing with storage deduplication for distributed storage. They proposed a novel plan in light of systems including polynomial based verification labels and homomorphic straight authenticators. Their plan permits deduplication of both documents and their corresponding confirmation labels. Information honesty evaluating and capacity deduplication are accomplished at the same time. Open evaluating and bunch inspecting are both bolstered. Yet, plausibility of supporting deduplication huge information was not talked about in this work. In order to diminish workloads because of copy documents, Wu et al. proposed File Name Servers (INS) to oversee not just record stockpiling, information deduplication, advanced hub determination, and server stack adjusting, additionally document compression, lump coordinating, genuine -time feedback control, IP data, and occupied level file checking. To oversee and upgrade stockpiling hubs in light of a customer - side transmission status by the proposed INS, all hubs must evoke ideal execution and offer appropriate re- sources to customers. In this way, not exclusively can the performance of a capacity framework be enhanced, yet the records can additionally be sensibly disseminated, diminishing the workload of the capacity hubs. Be that as it may, this work can't deduplifeline scrambled information.

Fan et al. proposed a mixture information deduplication mechanism that gives a commonsense arrangement incomplete semantic security. This arrangement underpins deduplication on plaintext and ciphertext. Be that as it may, this instrument can't bolster scrambled information deduplication extremely well. It works in light of the assumption that CSP knows the encryption key of information. In this manner it can't be utilized as a part of the circumstance that the CSP can't be completely trusted by the information holders or proprietors. In this paper, we apply ECC, PRE and symmetric encryption to deduplicate scrambled information. Our scheme can oppose the assaults said above in CE also, accomplish great execution with out keeping information holders on the web constantly. In the interim, it likewise guarantees the classification of put away information and backings advanced rights administration. We mean to accomplish deduplication on encoded huge information in cloud .

3 PROBLEM STATEMENTS

3.1 System and Security

We propose a plan to deduplicate encoded information at CSP by applying PRE to issue keys to various approved information holders in view of information possession challenge. It is relevant in situations where information holders are not accessible for deduplication control. The framework contains three sorts of elements: 1) CSP that offers stockpiling administrations and can't be completely trusted since it is interested aboutthe substance of put away information, yet ought to perform sincerely on information stockpiling with a specific end goal to increase business benefits; 2) information holder (ui) that transfers and spares its information at CSP. In the framework, it is conceivable to have various qualified information holders (ui,i=1,...,n)that could spare the same encoded crude information in CSP. The information holder that produces or makes the document is viewed as information proprietor. It has higher need than other typical information holders, which will be exhibited in Section 4; 3) an approved gathering (AP) that does not conspire with CSP and is completely trusted by the information holders to confirm information possession and handle information deduplication. For this situation, AP can't know the information put away in CSP and CSP ought not know the plain client information in its stockpiling. In principle it is conceivable that CPS and its clients (e.g., information holders) can plot. Practically speaking, notwithstanding, such plot could make the CSP lose notoriety because of information spillage. A negative effect of awful notoriety is the CSP will lose its clients lastly make it lose benefits. Then again, the CSP clients (e.g., information holders) could lose their comfort and advantages of putting away information in CSP because of awful notoriety of distributed storage administrations. In this manner, the arrangement amongst CSP and its clients is not productive for them two. Solid investigation in view of Game Theory is given in. Accordingly, we hold such a supposition as: CSP does not plot with its clients, e.g., performing re-encryption for unapproved clients to enable them to get to information. Extra presumptions include: information holders sincerely give the scrambled hash codes of information for possession confirmation. The information proprietor has the most elevated need. An information holder ought to give a legitimate endorsement so as to ask for a unique treatment. Clients, CSP and AP convey through a protected channel (e.g., SSL) with each other. CSP can validate its clients during the time spent cloud information storage .We additionally expect that the client arrangement (u) for information stockpiling, sharing and deduplication is given to CSP amid client enrollment.

3.2 Notation

The two major preliminaries here are Proxy Re-Encryption and Symmetric Encryption.

Table 1 . System Notation

Key	**Description**
(pki,ski)	The public key and secret keyof user ui for PRE.
DEKi	The symmetric key of ui.
H( )	The hash function.
CT	The cipertext
CK	The cipherkey
M	The user data
KG	The key generation algorithm of PRE
RG	The re-encryption key generation algorithm
R	The re-enccryption algorithm of PRE
D	The decryption algorithm of PRE
Encrypt (DEK, M)	The encryption function on M with DEK
Decrypt(DEK, CT)	The symmetric decryption function on CT with DEK

(V_i, s_i)	The key user ui in Eillptic Curve Cryptography(ECC) for duplication check
P	The base point in ECC
$X = H(H(M)*P)$	The token used for data duplication check

4 SCHEME

Our plan contains the accompanying principle angles Encrypted Data Upload: If information duplication check is negative, the information holder scrambles its information utilizing a radomly chose symmetric key DEK keeping in mind the end goal to guarantee the security and protection of information, and stores the scrambled information at CSP together with the token utilized for dataduplication check. The information holder scrambles DEK with pkAP and passes the scrambled key to CSP.

Information Deduplication: Data duplication happens when information holder u tries to store similar information that has been put away as of now at CSP. This is checked by CSP through token correlation. On the off chance that the examination is certain, CSP contacts AP for deduplication by giving the token and the information holder's PRE open key. The AP challenges information proprietorship, checks the qualification of the information holder, and after that issues a re-encryption key that can change over the scrambled DEK to a shape that must be unscrambled by the qualified information holder.

Information Deletion: When the information holder erases information from CSP, CSP right off the bat deals with the records of copied information holders by evacuating the duplication record of this client. On the off chance that the rest records are not vacant, the CSP won't erase the put away encoded information, yet piece information access from the holder that solicitations information cancellation. On the off chance that the rest records are void, the scrambled information ought to be evacuated at CSP.

Information Owner Management: on the off chance that that a genuine information proprietor transfers the information later than the information holder, the CSP can figure out how to spare the information scrambled by the genuine information proprietor at the cloud with the proprietor created DEK and later on, AP underpins re-encryption of DEK at CSP for qualified information holders.

Encoded Data Update: on the off chance that that DEK is refreshed by an information proprietor with DEK' and the new encoded crude information is given to CSP to supplant old stockpiling for the reason of accomplishing better security, CSP issues the new re-encoded DEK' to all information holders with the support of AP.

5 SECURITY ANALYSIS and PERFORMANCE EVALUATION

5.1 Security Analysis

Our plan Provides a safe way to deal with secure and deduplicate the information put away in cloud by hiding plaintext from both CSP and AP. The security of the expert postured plan is guaranteed by PRE hypothesis, symmetric key encryption and ECC hypothesis.

Recommendation 1.The participation of CSP and AP without agreement ensures that lone qualified clients can get to plain information M and the information can be deduplicate security. CSP knows CK encoded with pkAP, however AP does not share its own particular mystery key skAP with CSP. In this way CSP can't known DEK and afterward M. AP has no real way to get to M since its get to is obstructed by CSP in spite of the fact that AP could acquire DEK. What's more, we apply appropriate administration conventions to bolster information stockpiling administration and information proprietor administration to accomplish deduplication in the meantime.

Recommendation 2. H(M) that is the way to pass the duplication check can't be gotten by vindictive cloud clients. In light of the above security investigation, H(M)is not transmitted between the included gatherings and can't be gotten by whatever other gatherings. Notwithstanding when CSP plots with a noxious cloud client, the vindictive client can just get the token $H(H(M)\times P)$, not H(M), consequently difficult to get to M put away at cloud by passing the possession challenge raised by AP.

Recommendation 3.To pass the possession confirmation of AP, a cloud client should without a doubt have information M.

5.2 Computation Complexity

The proposed plot includes four sorts of framework parts: information proprietor, information holder, CSP and AP. To show the calculation unpredictability in points of interest, we receive AES for symmetric encryption, ECC and PRE proposed in [8]. We dissect the unpredictability of transferring one information document as be-low.

Information Owner: Regarded as the main information uploader, it is responsible for four operations: framework setup, information encryption, key encryption, and token $H(H(M)\times P)$generation. In setup, the key era of PRE incorporates 1 exponentiation. The ECC key era needs one point increase. What's more, framework setup takes once for all information stockpiling operations. The calculation many-sided quality of encoding information utilizing DEK relies on upon the span of information, which is inescapable in any cryptographic techniques for securing the information.

CSP: A client transfers its information to CSP by sending token $H(H(M)\times P)$. CSP ought to first check if a similar token has existed (by contrasting the token and the records in CSP, which is inescapable in any deduplication plans). At that point, CSP spares the information if the token does not exist. In the event that the information holder transfers similar information, CSP contacts AP for picking up a re-encryption key if the possession test is certain.

Information Holder: When information holder ui transfers similar information that has been put away in CSP, it creates token H(H(M)×P) as the information proprietor has done, which needs one point duplication.

AP: AP is in charge of the re-encryption key administration. It challenges information proprietorship by arbitrarily choosing c, decoding y and looking at H(yP+ cVi). We can see that the correspondence cost of our plan is light and it is not affected by the extent of transferred information. Consequently, the proposed plan is appropriate for supporting enormous information deduplication with respect to correspondence cost.

5.3 Further Discussions

The proposed conspire has the accompanying extra promotion vantages.

Adaptability: The proposed plan can adaptably bolster get to control on scrambled information with deduplication. One information holder can adaptably refresh DEK. The new key can be effortlessly issued to other information holders or qualified information clients by CSP with an ease, particularly when AP has issued the re-encyption key as of now. Information denial can be acknowledged by blocking information access at CSP and dismissing key re-encryption on a recently connected key DEK'. The point by point procedure of information denial is depicted in [34].

Minimal effort of capacity: The plan can clearly spare the storage room of CSP since it just stores one duplicate of similar information that is shared by information proprietor and information holders. Putting away deduplication records involves some stockpiling or memory for sparing token pki and xi (just 1024+160 bits). Be that as it may, contrasting and the enormous volume of copied information, this stockpiling expense can be disregarded.

Enormous information bolster: The proposed plan can effectively perform huge information deduplication. Initially, copied huge information transfer is productive in light of the fact that lone xi and pki are sent to CSP. CSP performs hash correlation and afterward contacts AP to test proprietorship for issuing a re-encryption key. The calculation and correspondence cost of this procedure (including possession challenge, re-encryption key era, CK re-encryption and re-encoded key decoding) is not impacted by the span of huge information. Second, transferring ciphertext CT is inescapable in all plans for deduplication. The proposed conspire just presents somewhat additional correspondence load (i.e.,CK) and a tad bit extra correspondence taken a toll for possession challenge. Contrasted and huge information transfer cost and capacity cost, they are exceptionally inconsequential and effective.

6. CONCLUSION

Overseeing encoded information with deduplication is vital and critical practically speaking for accomplishing an effective distributed storage benefit, especailly for enormous information stockpiling. In this paper, we proposed a down to earth plan to deal with the encoded enormous information in cloud with deduplication in view of proprietorship test and PRE. Our plan can adaptably bolster information refresh and imparting to deduplication notwithstanding when the information holders are disconnected. Encoded information can be safely gotten to in light of the fact that exclusive approved information holders can get the symmetric keys utilized for information unscrambling. Broad execution investigation and test demonstrated that our plan is secure and productive under the de-scribed security show and extremely reasonable for huge information deduplication. The aftereffects of our PC recreations additionally demonstrated the practicability of our plan. Future work incorporates streamlining our outline and execution for down to earth organization and concentrate erifiable calculation to guarantee that CSP carries on of course in deduplication administration.

REFERENCES

[1] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski,G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, “A viewof cloud computing,”Communication of the ACM, vol. 53, no. 4, pp.50–58, 2010.

[2] J. Yuan and S. Yu, “Secure and constant cost public cloud storage auditing with deduplication,” invitee Conference on Communications and Network Security (CNS), 2013, pp. 145–153.

[3] S. Halevi, D. Harnik, B. Pinkas, and A. Shulman-Peleg, “Proofs ofownership in remote storage systems,” in Proceedings of the 18th ACMConference on Computer and Communications Security. ACM, 2011,pp. 491–500.

[4] S. Keelveedhi, M. Bellare, and T. Ristenpart, “Dupless: Server-aided encryption for deduplicated storage,” in Proceedings of the22Nd USENIX Conference on Security, ser. SEC’13. Washington,D.C.: USENIX Association, 2013,pp.179194.[Online].Available:https://www.usenix.org/conference/usenixsecurity13/technical-essions/presentation/bellare

[5] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson,and D. Song, “Provable data possession at untrusted stores,” inProceed-ings of the 14th ACM Conference on Computer and CommunicationsSecurity, ser. CCS ’07. New York, NY, USA: ACM, 2007, pp. 598–609.

[6] G. Ateniese, R. Burns, R. Curtmola, J. Herring, O. Khan, L. Kissner,Z. Peterson, and D. Song, “Remote data checking using provable datapossession,”ACM Trans. Inf. Syst. Secur., vol. 14, no. 1, pp. 12:1–12:34,2011.

[7] G. Ateniese, R. Di Pietro, L. V. Mancini, and G. Tsudik, “Scalableand efficient provable data possession,” in Proceedings of the 4th International Conference on Security and Privacy in CommunicationNetowrks, ser.

SecureComm ’08. New York, NY, USA: ACM, 2008,pp. 9:1–9:10.

[8] C. Erway, A. K¨upc¸¨u, C. Papamanthou, and R. Tamassia, “Dynamic provable data possession,” in Proceedings of the 16th ACM Conference on Computer and Communications Security, ser. CCS ’09. New York,NY, USA: ACM, 2009, pp. 213–222.

[9] F. Seb´e, J. Domingo-Ferrer, A. Martinez-Balleste, Y. Deswarte, andJ.-J. Quisquater, “Efficient remote data possession checking in critical information infrastructures,” IEEE Trans. on Knowl. and Data Eng.,vol. 20, no. 8, pp. 1034–1038, 2008.

[10] H. Wang, “Proxy provable data possession in public clouds,” IEEE Transactions on Services Computing, vol. 6, no. 4, pp. 551–559, 2013.

[11] Y. Zhu, H. Hu, G.-J. Ahn, and M. Yu, “Cooperative provable data possession for integrity verification in multicolor storage,” IEEE Trans-actions on Parallel and Distributed Systems, vol. 23, no. 12, pp. 2231–2244, 2012.

[12] H. Shacham and B. Waters, “Compact proofs of retrievability,” in Proceedings of the 14th International Conference on the Theory and Application of Cryptology and Information Security: Advances in Cryp-tology, ser. ASIACRYPT ’08. Springer Berlin Heidelberg, 2008, pp.90–107.

[13] Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, “Enabling public verifiability and data dynamics for storage security in cloud computing, “in Computer Security – ESORICS 2009, M. Backes and P. Ning, Eds.,vol. 5789. Springer Berlin Heidelberg, 2009, pp. 355–370.

[14] J. Xu and E.-C. Chang, “Towards efficient proofs of retrievability,” in Proceedings of the 7th ACM Symposium on Information, Computer and Communications Security, ser. ASIACCS ’12. New York, NY, USA:ACM, 2012, pp. 79–80.

[15] E. Stefanov, M. van Dijk, A. Juels, and A. Oprea, “Iris: A scalablecloud file system with efficient integrity checks,” inProceedings of the28th Annual Computer Security Applications Conference, ser. ACSAC’12. New York, NY, USA: ACM, 2012, pp. 229–238.

[16] M. Azraoui, K. Elkhiyaoui, R. Molva, and M.¨Onen, “Stealthguard:Proofs of retrievability with hidden watchdogs,” in Computer Security -ESORICS 2014, ser. Lecture Notes in Computer Science, M. Kutyłowskiand J. Vaidya, Eds., vol. 8712. Springer International Publishing, 2014,pp. 239–256.

[17] J. Li, X. Tan, X. Chen, and D. Wong, “An efficient proof of retrievability with public auditing in cloud computing,” in5th International Con-ference on Intelligent Networking and Collaborative Systems (INCoS),2013, pp. 93–98.

[18] J. Li, X. Chen, M. Li, J. Li, P. Lee, and W. Lou, “Secure deduplication with efficient and reliable convergent key management,” IEEE Transactions on Parallel and Distributed Systems, vol. 25, no. 6, pp.1615–1625, June 2014.

[19] R. Di Pietro and A. Sorniotti, “Boosting efficiency and security in proof of ownership for deduplication,” inProceedings of the 7th ACM Symposium on Information, Computer and Communications Security,ser. ASIACCS ’12. New York, NY, USA: ACM, 2012, pp. 81–82.

[20] W. K. Ng, Y. Wen, and H. Zhu, “Private data deduplication protocolsin cloud storage,” in Proceedings of the 27th Annual ACM Symposiumon Applied Computing, ser. SAC ’12. New York, NY, USA: ACM,2012, pp. 441–446.

[21] J. Douceur, A. Adya, W. Bolosky, P. Simon, and M. Theimer, “Reclaim-ing space from duplicate files in a serverless distributed file system,”in 22nd International Conference on Distributed Computing Systems,2002, pp. 617–624.

[22] M. Bellare, S. Keelveedhi, and T. Ristenpart, “Message-locked encryp-tion and secure deduplication,” in Advances in Cryptology – EURO-CRYPT 2013, ser. Lecture Notes in Computer Science, T. Johanssonand P. Nguyen, Eds. Springer Berlin Heidelberg, 2013, vol. 7881, pp.296–312.

[23] M. Abadi, D. Boneh, I. Mironov, A. Raghunathan, and G. Segev,“Message-locked encryption for lock-dependent messages,” inAdvancesin Cryptology – CRYPTO 2013, ser. Lecture Notes in Computer Science,R. Canetti and J. Garay, Eds. Springer Berlin Heidelberg, 2013, vol.8042, pp. 374–391.