GLOBAL JOURNAL OF ENGINEERING SCIENCE AND RESEARCHES

ELIMINATING DUPLICATE DATA WITH POS FOR MULTI USER ENVIRONMENT

Bhagawat Pooja N.[*1] **,Bhagawat Pallavi S.**[2] **& Bhor Ashwini D**[3]
[*1,2&3]BE, Department Of Computer Engineering

ABSTRACT

Dynamic Proof of Storage (PoS) is a useful cryptographic primitive that enablesa user to check the integrity of outsourced less and to efficiently updatethe files in a cloud server. Although researchers have proposed many dynamicPoS schemes in single user environments, the problem in multi-user environmentshas not been investigated sufficiently. A practical multi-user cloudstorage system needs the secure client-side cross-user deduplication technique,which allows a user to skip the uploading process and obtain the ownership ofthe files immediately, when other owners of the same files have uploaded themto the cloud server. To the best of our knowledge, none of the existing dynamicPoSs can support this technique. In this system, we introduce the concept ofdeduplicatable dynamic proof of storage and propose an efficient constructioncalled DeyPoS, to achieve dynamic PoS and secure cross-user deduplication,simultaneously. Considering this challenges of structure diversity and privatetag generation. We will prove the security of our construction, and the theoreticalanalysis and experimental results which will show that our construction isefficient in practice.

Keywords: *Deduplication ,Proof of ownership, Dynamic proof of storage, Cloud Computing.*

I. INTRODUCTION

Storage outsourcing is turning into additional and additional enticing to each trade and tutorial because of the benefits of low value, high accessibility, and straightforward sharing. Collectively of the storage outsourcing forms, cloud storage gains wide attention in recent years. Several firms, like Amazon, Google, and Microsoft, give their own cloud storage services, wherever users will transfer their files to the servers, access them from varied devices, and share them with the others. Though cloud storage services are wide adopted in current days, there still stay several security problems and potential threats .Data integrity is one among the foremost vital properties once a user outsources its files to cloud storage. Users ought to be convinced that the files keep within the server don't seem to be tampered. Ancient techniques for safeguarding knowledge integrity, like message authentication codes (MACs) and digital signatures, need users to transfer all of the files from the cloud server for verification, that incurs a significant communication value. These techniques don't seem to be appropriate for cloud storage services wherever users could check the integrity oftentimes, like each hour. Thus, researchers introduced Proof of Storage (PoS) for checking the integrity while not downloading files from the cloud server. what is more, users may need many dynamic operations, like modification, insertion, and deletion, to update their files, whereas maintaining the potential of PoS. Dynamic PoS is projected for such dynamic operations. In distinction with PoS, dynamic PoS employs structures, like the Merkle tree. Thus, once dynamic operations are dead, users regenerate tags (which are used for integrity checking, like MACs and signatures) for the updated blocks solely, rather than create for all blocks. to raised perceive the subsequent contents, we tend to gift additional details concerning PoS and dynamic PoS. In these schemes, every block of a file is hooked up a(cryptographic) tag that is employed for substantiating the integrity of that block. once a champion desires to ascertain the integrity of a file, it every which way selects some block indexes of the file, and sends them to the cloud server. Consistent with these challenged indexes, the cloud server returns the corresponding blocks beside their tags. The champion checks the block integrity and index correctness. the previous are often directly bonded by cryptanalytic tags. a way to affect the latter is that the major distinction between PoS and dynamic PoS In most of the PoS schemes, the block index is "encoded" into its tag, which implies the champion will check the block integrity and index correctness at the same time. However, dynamic PoS cannot cypher the block indexes into tags, since the dynamic operations could modification several indexes of non-updated blocks, that incurs reserve computation and communication value. As an example, there's a file consisting of one thousand blocks, and a replacement block is inserted behind the second block of the file. Then,

998 block indexes of the first file are modified, which implies the user should generate and send 999 tags for this update. Structures are introduced in dynamic PoSs to unravel this challenge. As a result, the tags are hooked up to the structure instead of the block indexes .However, dynamic PoS remains to be improved in an exceedingly multi-user atmosphere, because of the necessity of cross-user American state duplication on the client-side. This means that users will skip the uploading method and acquire the possession of files now, as long because the uploaded files exist already within the cloud server. This method will cut back space for storing for the cloud server, and save transmission information measure for users. To the simplest of our data, There's no dynamic PoS that may support secure cross-user American state duplication.

II.   METHOD & MATERIAL
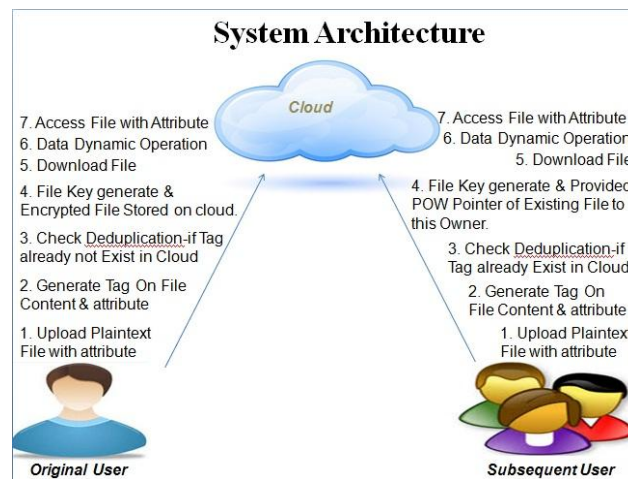
1. Pre-process Phase

Users will transfer their native files. The cloud server decides whether or not these files ought to be uploaded. If the transfer method is granted, enter the transfer phase; otherwise, enter the deduplication part.

2. Upload Phase

In the transfer part, the files to be uploaded don't exist within the cloud server. The first users encodes the native files and transfer them to the cloud server .

3. Deduplication Phase

The files to be uploaded exist already within the cloud server. the next users possess the files domestically and also the cloud server stores the structures of the files. ulterior users got to persuade the cloud server that they own the files while not uploading them to the cloud server. If these 3 phases (pre-process, upload, and deduplication) square measure dead just one occasion within the life cycle of a file from the angle of users. That is, these 3 phases seem only if users will transfer files. If these phases terminate unremarkably, i.e., users end transfering within the upload part, or they pass the verification within the deduplication part, we are saying that the users have the ownerships of the files.

4. Update Phase

Users could modify, insert, or delete some blocks of the files. Then, they update the corresponding components of the encoded files and also the structures within the cloud server, even the first files weren't uploaded by themselves.

Note that, users will update the files provided that they need the ownerships of the files, which suggests that the users ought to transfer the files within the transfer part or pass the verification within the deduplication.For each update, the cloud server needs to reserve the first file and also the structure if there exist different homeowners, and.record the updated a part of the file and also the structure. this permits users to update a file at the same time in our model, since every update is barely "attached" to the first file and structure.

5. Proof of Storage part

Users solely possess alittle constant size information domestically and that they need to examine whether or not the files square measure dependably hold on within the cloud server while not downloading them. The files might not be uploaded by these users, however they pass the deduplication part and prove that they need the ownerships of the files. Note that, the update part and also the proof of storage part will be dead multiple times within the life cycle of a file. Once the possession is verified, the users will randomly enter the update part and also the proof of storage part while not keeping the first files domestically.

III. LITRATURE SURVEY

i.   Title :A Secure and Dynamic Multi-keyword Ranked Search Scheme over Encrypted Cloud Data

Authors: Zhihua Xia, Xingming Sun, Qian Wang

In this paper, a secure, economical and dynamic search mechanism is projected, that supports not solely the correct multi-keyword hierarchical search however conjointly the dynamic deletion and insertion of documents. We have a tendency to construct a special keyword balanced binary tree because the index, and propose a "Greedy Depth-first Search" algorithmic program to get higher potency than linear search. Additionally, the parallel search process is administered to additional scale back the time price. the safety of the theme is protected against 2 threat models by exploitation the secure kNN algorithmic program. Experimental results demonstrate the potency of our projected theme. There is a unit still several challenge issues in radial SE schemes. Within the projected theme, owner is chargeable for generating change information and causation them to the cloud server.

ii.   Title: Security and Privacy in Cloud Computing: A Survey

Authors: Minqi Zhou, Rong Zhang, Wei Xie, Weining Qian, Aoying Zhou

Cloud Computing becomes a buzzword nowadays. More and more companies step into Cloud and provide services above on it. However, security and privacy issues impose strong barrier for users' adoption of Cloud systems and Cloud services. We observed the security and privacy concerns presented by an amount of Cloud Computing system providers in this paper. Nevertheless, those concerns are not adequate. More security strategies should be deployed in theCloud environment to achieve the 5 goals (i.e. availability, confidentiality, data integrity, control and audit) as well as privacy acts should be changed to adapt a new relationship between users and providers in the Cloud literature.

iii.   Title:From Security to Assurance in the Cloud: A Survey .

Authors: Claudio ardagna , Rasool asal.

Cloud tenants will use cloud resources at lower costs, and better performance and adaptability, than ancient on-premises resources, while not having to worry concerning infrastructure management. Still, cloud tenants stay involved with the cloud's level of service and therefore the nonfunctional properties their applications will judge. Within the previous couple of years, the analysis community have been specializing in the nonfunctional aspects of the cloud paradigm, among which cloud security stands out. Many approaches to security are delineate and summarized generally surveys on cloud security techniques. The survey during this article focuses on the interface

63

(C)*Global Journal Of Engineering Science And Researches*

This paper was presented in Technophilia 2017 at Jaihind Polytechnic, A/P. Kuran, Tal. Junnar, Dist. Pune

between cloud security and cloud security assurance. First, we offer an summary of the state of the art on cloud security. Then, we have a tendency to introduce the notion of cloud security assurance and analyze its growing impact on cloud security approaches. Finally, we have a tendency to gift some recommendations for the event of next-generation cloud security and assurance solutions.

iv.   Hybrid Provable Data Possession at Untrusted Stores in Cloud Computing

Authors: Narn - Yih Lee,Yun - Kuan Chang

We centered the core problems, if Associate in Nursing untrusted server to store client data. we will demonstrable information possession within the model, that scale back the info block access, however conjointly scale back the number of computation on the server and shopper and server traffic. Our style and development on the PDP program is especially supported the usage of symmetrical and uneven cryptography system. It exceeds what we have a tendency to die within the past, the advance has delivered to the information measure, computation and storage system. And it applied the general public (third party) verification. Finally, we have a tendency to conjointly expect our program, it supports dynamic outsourcing of data build it a additional realistic application of cloud computing atmosphere.

IV. RESULT & DISCUSSION

1.   Pre-Process Phase

$e \leftarrow H(F)$, $id \leftarrow H(e)$.

Where,

id = File Identity.

2.   Upload Phase

File $F = (m1, . . . ,mn)$.

The user first invokes the encoding according,

$(C, T) \leftarrow Encode(e, F)$

Where,

$m1, . . . ,mn$= Represents I^{th} block of file.

e = Encryption key.

3.   The Deduplication Phase

If a file announced by a user in the pre-process phase exists in the cloud server, the user goes into the deduplication phase and runs the deduplication protocol

$res \in \{0, 1\} \leftarrow Deduplicate\{U(e, F), S(T)\}$

Where,

res = Current uploading file.

e = Encryption Key.

F= Uploaded File.

The Update Phase

In this phase, a user can arbitrarily update the file, by invoking the update protocol

res ∈ {he∗, (C∗, T ∗)i,⊥} ← Update{U(e, ι,m,OP), S(C, T)}

Where,

res= Current updating file.

S(C,T)= Represent block to be uploaded.

The Proof of Storage Phase

At any time, users can go into the proof of storage phase if they have the ownerships of the files. The users and the cloud server run the checking protocol

res ∈ {0, 1} ←Check{S(C, T), U(e)}

Where,

res =Current file.

S(C,T)= Block of file.

V. CONCLUSION

We proposed the comprehensive requirements in multi-user cloud storage systems and introduced the model of deduplicatable dynamic PoS.We proposed the deduplicatable dynamic PoS and provide its security . The theoretical and experimental results which will show that our DeyPoS implementation is efficient, especially when the file size and the number of the challenged blocks are large.

VI. ACKNOWLEDGEMENTS

REFERENCES

1. *Z. Xia, X. Wang, X. Sun, and Q. Wang, "A Secure and Dynamic Multi-Keyword Ranked Search Scheme over Encrypted Cloud Data," IEEE Transactions on Parallel and Distributed Systems, vol. 27, no. 2, pp. 340–352, 2016.*
2. *Z. Xiao and Y. Xiao, "Security and privacy in cloud computing," IEEE Communications Surveys Tutorials, vol. 15, no. 2, pp. 843859, 2013.*

3. *C. A. Ardagna, R. Asal, E. Damiani, and Q. H. Vu, "From Security to Assurance in the Cloud: A Survey," ACM Comput. Surv., vol. 48, no. 1, pp. 2:1–2:50, 2015.*
4. *G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable data possession at untrusted stores," in Proc. of CCS, pp. 598–609, 2007.*
5. *G. Ateniese, R. Di Pietro, L. V. Mancini, and G. Tsudik, "Scalable and Efficient Provable Data Possession," in Proc. Of SecureComm, pp. 1–10, 2008. G. Ateniese, S. Kamara, and J. Katz, "Proofs of storage from homomorphic identification protocols," in Proc. of ASIACRYPT, pp. 319–333, 2009.*
6. *C. Erway, A. K¨upc¨u, C. Papamanthou, and R. Tamassia, "Dynamic provable data possession," in Proc. of CCS, pp. 213–222, 2009.*
7. *R. Tamassia, "Authenticated Data Structures," in Proc. of ESA, pp. 2–5, 2003.*
8. *Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, "Enabling public verifiability and data dynamics for storage security in cloud computing," in Proc. of ESORICS, pp. 355–370, 2009.*
9. *F. Armknecht, J.-M. Bohli, G. O. Karame, Z. Liu, and C. A. Reuter, "Outsourced proofs of retrievability," in Proc. of CCS, pp. 831–843, 2014.*
10. *H. Shacham and B. Waters, "Compact Proofs of Retrievability," Journal of Cryptology, vol. 26, no. 3, pp. 442–483, 2013*

(C)*Global Journal Of Engineering Science And Researches*

This paper was presented in Technophilia 2017 at Jaihind Polytechnic, A/P. Kuran, Tal. Junnar, Dist. Pune