Dynamic Proof of Storage for Avoiding Deduplication in Multiple User Environment

Bhor Ashwini D, Bhagawat Pallavi S, Bhagawat Pooja N, Prof .Gunjal S.D.

***Abstract*— Dynamic Proof of Storage (PoS) is a useful cryptographic primitive that enables a user to check the integrity of outsourced less and to efficiently update the files in a cloud server. Although researchers have proposed many dynamic PoS schemes in single user environments, the problem in multi-user environments has not been investigated adequately. A practical multi-user cloud storage system needs the secure client-side cross-user deduplication technique, which allows a user to skip the uploading process and obtain the rights of the files immediately, when other owners of the same files have uploaded them to the cloud server. To the best of our knowledge, none of the existing dynamic PoSs can support this technique. In this system, we introduce the concept of deduplicatable dynamic proof of storage and propose an efficient construction called DeyPoS, to achieve dynamic PoS and secure cross-user deduplication, simultaneously. Considering this challenges of structure diversity and private tag generation. We will prove the security of our construction, and the theoretical analysis and experimental results which will show that our construction is efficient in practice.**

***Index Terms*— Deduplication, Proof of ownership, Cloud Computing**

I. INTRODUCTION

Stroage outsourcing is becoming more and more attractive to both industry and academic due to the advantages of low cost, high accessibility, and easy sharing. Many companies, such as Amazon, Google, and Microsoft, provide their own cloud storage services, where users can upload their files to the servers, access them from various devices, and share them with the others. Although cloud storage services are widely adopted in current days, there still remain many security issues and potential threats[1][7].Data integrity is one of the most important properties when a user outsources its files to cloud storage. Users should be convinced that the files stored in the server are not tampered. Thus, researchers introduced Proof of Storage (PoS)[3] for checking the integrity without downloading files from the cloud server. Furthermore, users may also require a few dynamic operations, such as modification, insertion, and deletion, to update their files, while maintaining the abilities of PoS. Thus, when dynamic operations are executed, users regenerate tags for the updated blocks only, instead of regenerating for all blocks. In these system, each block of a file is attached a(cryptographic) tag which is used for verifying the integrity of that block. When a verifier wants to check the integrity of a file, it randomly selects some block indexes of the file, and sends them to the cloud server. According to these challenged indexes[9], the cloud server returns the corresponding blocks along with their tags. The verifier checks the block integrity and index correctness. The former can be directly guaranteed by cryptographic tags. However, dynamic PoS [4][6]cannot encode the block indexes into tags, since. However, dynamic PoS remains to be improved in a multi-user environment, due to the requirement of cross-user deduplication on the client-side. This indicates that users can skip the uploading process and obtain the ownership of files immediately, as long as the uploaded files already exist in the cloud server. This technique can reduce storage space for the cloud server, and save transmission bandwidth for users. To the best of our knowledge, there is no dynamic PoS that can support secure cross-user deduplication.

II. LITRATURE SURVEY

1] 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[2].

2]Title: *Security and Privacy in Cloud Computing.*
Authors: Minqi Zhou, Rong Zhang, Wei Xie, Weining Qian, Aoying Zhou
Cloud Computing becomes a exhortation 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 the Cloud 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[1].

3]Title: *From Security to Assurance in the Cloud.*
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 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[7].

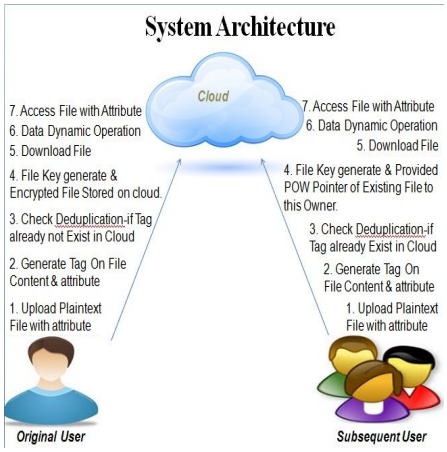
4] *Hybrid Provable Data Possession at Untrusted Stores in Cloud Computing.*
Authors: Narn - Yih Lee, Yun - Kuan Chang.
We centered the core problems, if Associate in Nursing untrusted server to store client data. we will demonstrable information possession within the model, that scale back the info 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[5].

III. EXISTING SYSTEM

Despite the fact that deduplication strategy can spare the storage for the distributed storage service provider, it decreases the unwavering quality of the system.
Deduplication system and distributed storage system are expected by clients and applications for higher unwavering quality, particularly in archival storage system where information are basic and ought to be saved over long time periods.

IV. PRAPOSED SYSTEM

No Such system of Dynamic proof of storage will Achive cross user deduplication. To remove this drawbacks we impliment Deduplicatable dynamic proof of storage.



As shown in Fig. for every file, original user is that the user World Health Organization uploaded the file to the cloud server, whereas ulterior user is that the user World Health Organization established the possession of the file however didn t truly transfer the file to the cloud server. There square measure 5 phases during a deduplicatable dynamic PoS system: pre-process, upload, deduplication, update, and proof of storage.

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[10] .

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[10].

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[10].

5. Proof of Storage part
Users solely possess a little 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[10].

V. MATHEMATICAL CALCULATION

1.Pre-Process Phase
e ← H(F), id ← H(e).
Where,
id = File Identity.

2.Upload Phase
File F = (m1, . . . ,mn).
The user first invokes the encoding according,
(C, T) ← Encode(e, F)
Where,
m1, . . . ,mn= Represents Ith 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 ∈ {0, 1} ← Deduplicate{U(e, F), S(T)}
Where,
res = Current uploading file.
e = Encryption Key.
F= Uploaded File.

4. 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.

5. 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.

VI. ALGORITHM

1.MD5
MD5 algorithm takes input message of arbitrary length and generates 128-bit long output hash.In our system the Md5 algorithm is use for checking the deduplication by using creating hash value.Every file content is create the different hash value and according to that it avoid the storage space. MD5 hash algorithm consist of 5 steps :

Step 1. Append Padding Bits

Step 2. Append Length

Step 3. Initialize MD Buffer

Step 4. Process Message in 16-Word Blocks

Step 5. Output

2.AES

The encryption process uses a set of specially derived keys called round keys. These are applied, along with other operations, on an array of data that holds exactly one block of data the data to be encrypted. This array we call the state array:

1.Derive the set of round keys from the cipher key.

2.Initialize the state array with the block data (plaintext).

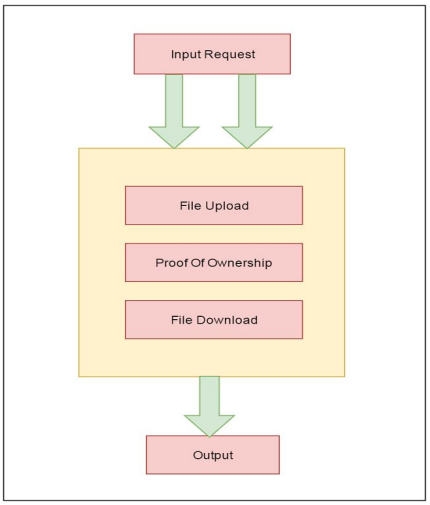
3.Add the initial round key to the starting state array.

4.Perform nine rounds of state manipulation.

5.Perform the tenth and final round of state manipulation.

6.Copy the final state array out as the encrypted data (ciphertext).

VII. DESIGHN AND IMPLEMENTATION



VIII. FUTURE SCOPE

A.It is usable in Social networking sites or applications using cloud and handles many users and uploading large amount of same data in cloud.

B.DeyPoS mechanism helps to manage all data on cloud without creating duplicate copies of files of various sites.

C.It also helps to provide access or grant ownership permissions to site users.

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.

ACKNOWLEDGMENT

We might want to thank the analysts and also distributers for making their assets accessible. We additionally appreciative to commentator for their significant recommendations furthermore thank the school powers for giving the obliged base and backing

REFERENCES

[1]Z. Xiao and Y. Xiao, "Security and privacy in cloud computing," *IEEE Communications Surveys Tutorials*, vol. 15, no. 2, pp. 843859, 2013.

[2] 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.

[3] G. Ateniese, S. Kamara, and J. Katz, "Proofs of storage from homomorphic identification protocols," in Proc. of ASIACRYPT, pp. 319–333, 2009.

[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] Narn - Yih Tai,Taiwan, Yun - Kuan Chang"Hybrid Provable Data Possession at Untrusted Stores In Cloud Computing"in 2011 IEEE 17th International Conference on Parallel and Distributed Systems.

[6] 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.

[7] C. Erway, A. K¨upc ¨u, C. Papamanthou, and R. Tamassia, "Dynamic provable data possession," in *Proc. of CCS*, pp. 213–222, 2009.

[8] 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.

[9] 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.

[10] Kun He, Jing Chen, Ruiying Du, Qianhong Wu, Guoliang Xue, and Xiang Zhang "DeyPoS: Deduplicatable Dynamic Proof of Storage for Multi-User Environments" DOI 10.1109/TC.2016.2560812, IEEE