USING BLOCKCHAIN EXPANSION TECHNOLOGY TO AUDIT DATA INTEGRITY

[1]Y RAVI KUMAR, [2]N ANUSHA VENKAT

[1]Professor, Dept. of CSE, KMM Institute of Technology & Science, Tirupati.

[2]PG Student, Dept. of CSE, KMM Institute of Technology & Science, Tirupati.

ABSTRACT_ Disregarding the way that data decency is a critical issue, a rising number of clients are re-appropriating data to the cloud. Because of its decentralization and permanence, specialists are progressively going to blockchain to supplant outsider examiners. To address the significant costs of keeping up with the blockchain network and allowing clients to make new blocks due to the rapid development of blocks in the current blockchain innovation's information honesty review conspire, this paper proposes an information respectability framework in light of blockchain development innovation. On the chief chain and its subchains, smart arrangements are done by clients and cloud expert centers (CSP). The sub-chain sends the results of its extensive and frequent computations to the main chain on a regular basis or as needed to ensure their finality. To do whatever it takes not to impact the client experience due to correspondence with the CSP during the audit, the possibility of non-astute survey is introduced. In order to guarantee the confidentiality of data, a prize pool instrument is used. A comprehensive examination from a variety of perspectives, including capacity, cluster evaluation, and information consistency, demonstrates the plan's accuracy. The Ethereum blockchain stage's tests demonstrate how this strategy can effectively reduce limit and computational overhead..

1.INTRODUCTION

In the present essential [31] and reinforcement [26], [39], and [42] capacity frameworks, Piece BASED deduplication is much of the time used to save a ton of room. It refers to all copy lumps as the actual duplicate via small size references, but it only stores one actual copy of copy pieces. It has been demonstrated that deduplication effectively reduces the storage space required by primary storage by 50% [31] and backup storage by up to 98 percent [39]. This prods the wide association of deduplication in various business dispersed capacity organizations (e.g., Dropbox, Google Drive, Bitcasa, Mozy, and Memopal) to lessen huge limit costs [18]. Scrambled deduplication adds an encryption layer to deduplication [7, 8] to provide classification assurances. Prior to being composed to deduplicated capacity, each piece is deterministically encoded using symmetric-key encryption with a key obtained from the bulk content (for example, the key is set to be the cryptographic hash of piece content [14]). Along these lines, we can apply deduplication to the scrambled pieces to save space since copy lumps actually have a similar substance even after encryption. Numerous studies (e.g., [5], [7], [25], [33], and [36]) have developed a variety of encrypted deduplication schemes for the purpose of efficiently managing outsourced data in cloud storage. In addition to storing data that is not duplicated, a deduplicated storage system must also store deduplication metadata. Metadata deduplication can be broken down into two groups. The framework keeps a finger impression list that tracks the fingerprints of all lumps that have previously been put away to check whether they are indistinguishable. Moreover, to allow a record to be duplicated, the system keeps a report recipe that holds the mappings from the knots in the record to the references of the relating physical copies.\

Deduplication metadata is broadly known to cause high limit above [11], [21], [30], especially for the incredibly tedious obligations (e.g., fortifications) as the metadata accumulating above ends up being more prevalent. We contend in this work that scrambled deduplication keeps key metadata, for example, key recipes that monitor the piece to-key mappings that make it conceivable to unscramble individual records, which brings about much higher metadata stockpiling above. Because they contain sensitive key information, key recipes must be managed separately from file recipes, encrypted using the master keys of file owners, and stored separately for each file owner. In real sending, encoded deduplication's stockpiling productivity might be undermined by a high metadata stockpiling above.

2. LITERATURE SURVEY

1.H. Wang, D. He, A. Fu, Q. Li, and Q. Wang, ‘‘Provable data possession with outsourced data transfer,’’ IEEE Trans. Services Comput., vol. 14, no. 6, pp. 1929–1939, Nov. 2021.

With the rapid development of cloud computing, more and more enterprises would like to upload and store their data in the public cloud. When the parts of the business of an enterprise are purchased by another enterprise, the corresponding data will be transferred to the acquiring enterprise. For the usual case, how to outsource the computation cost of data transfer to the cloud? How to ensure the remote purchased data integrity? Thus, it is important to study provable data possession with outsourced data transfer (DT-PDP). In this paper, for the first time, we propose the novel concept: DT-PDP. By taking use of DT-PDP, the following three security requirements can be satisfied: (1) the other un-purchased data security of acquired enterprise can be ensured; (2) the purchased data integrity and privacy can be ensured; (3) the data transferability’s computation can be outsourced to the public cloud servers. For the security concept of DT-PDP, we give its motivation, system model and security model. Then, we design a concrete DT-PDP scheme based on the bilinear pairings. At last, we analyze the security, efficiency and flexibility of the concrete DT-PDP scheme. It shows that our scheme is provably secure and efficient.

2. J. Chang, B. Shao, Y. Ji, M. Xu, and R. Xue, ‘‘Secure network coding from secure proof of retrievability,’’ Sci. China Inf. Sci., vol. 64, no. 12, Dec. 2021, Art. no. 229301

In recent years, storage-as-a-service has emerged as a commercial alternative for user's local data storage due to its features include less initial infrastructure setup, relief from maintenance overhead, and universal access to the data irrespective of the location and devices [5]. However, it also faces several security threats. One of the most serious threats is the integrity of user's stored data. In particular, when storing the data file to a cloud service provider (CSP), a user (or data owner) will delete it from his/her local devices and hence lose local control of it. In this case, CSP may discard some user's rarely accessed data to save its space and earn more profit. Meanwhile, the CSP can lie about the fact. Obviously, it is extremely unfavorable for users. Proof of retrievability (PoR) protocol is just one of initial attempts to formulize the notion of "remotely and reliably checking data's integrity without downloading the whole data file".

3.N. Döttling and S. Garg, ''Identity-based encryption from the Diffie–Hellman assumption,'' J. ACM, vol. 68, no. 3, pp. 1–46, Mar. 2021

We provide the first constructions of identity-based encryption and hierarchical identitybased encryption based on the hardness of the (Computational) Diffie-Hellman Problem (without use of groups with pairings) or Factoring. Our construction achieves the standard notion of identity-based encryption as considered by Boneh and Franklin [CRYPTO 2001]. We bypass known impossibility results using garbled circuits that make a non-black-box use of the underlying cryptographic primitives.

4. H. Yan, J. Li, and Y. Zhang, ''Remote data checking with a designated verifier in cloud storage,'' IEEE Syst. J., vol. 14, no. 2, pp. 1788–1797, Jun. 2020.

Remote data possession checking (RDPC) supplies an efficient manner to verify the integrity of the files stored in cloud storage. Public verification allows anyone to check the integrity of remote data so that it has a wider application in public cloud storage. Private verification just allows the data owner to verify the data integrity, which is mainly applied for the verification of secret data. However, in many real applications, the data owner expects a specific user to check the files in cloud storage, whereas others cannot execute such work. It is obvious that neither public verification nor private verification can satisfy such a requirement. To solve this issue, Ren et al. provided a designated-verifier provable data possession (DV-PDP) protocol. Unfortunately, the DV-PDP is insecure against replay attack launched by the malicious cloud server. To overcome this shortcoming, we present a new RDPC scheme with the designated verifier, in which the data owner specifies a unique verifier to check the data integrity. Based on the computational Diffie-Hellman assumption, we prove the security for our RDPC scheme in a random oracle model. The theoretical analysis and experiment results indicate that our scheme has less communication, storage, and computation overhead while achieving high error detection probability.

5. J. Chang, H. Wang, F. Wang, A. Zhang, and Y. Ji, ‘‘RKA security for identity-based signature scheme,’’ IEEE Access, vol. 8, pp. 17833–17841, 2020

Related-key attack (RKA) is a kind of side-channel attack considered for kinds of cryptographic primitives, such as public key encryption, digital signature, pseudorandom functions etc. However, we note that the RKA-security seems to be not considered for identity-based signature (IBS), which is an important primitive for identity-based cryptography and proposed by Shamir in 1984. In this paper, for the first time, we introduce the RKA security into IBS schemes and try to define the security model for it. More specifically, we consider the RKA occurs in the users' signing key or the master key of the key-generation center (KGC), which derives two kinds of RKA securities for IBS. Meanwhile, we illustrate that the most efficient Schnorr-like IBS scheme proposed by Galindo and Garcia is RKA-insecure by launching a simple RKA. However, a slight modification of it yields a RKA-secure IBS scheme, for which we give the detailed security proof in the random oracle. Finally, the performance analysis shows that the modified scheme is still extremely efficient but has higher security

3.PROPOSED SYSTEM

1) The suggested framework includes a proposal for an information uprightness review convention in the context of plasma brilliant agreements. By implementing plasma sub-chains, smart contracts, and smart contracts on the main chain and sub-chains, this protocol has the ability to lessen the storage burden on the main chain and slow down the growth rate.

2) For the suggested system, a batch auditing technique is suggested. Multiple audit jobs can be batch processed at once using this approach. Executing the TPA audit protocol requires nothing in the way of computational and communication overhead. The concept of non-intuitive review is put forth in an effort to limit how much the client experience will be impacted by communication with the CSP during the review cycle. The award pool mechanism is used to assure the accuracy of the review, and the confirmation hub can receive appropriate rewards.

3) In the suggested framework, an analysis of the plan's security shows how it might achieve the common security objectives. Multiple ether block chain trials also demonstrated the effectiveness and viability of the scheme.

3.1 IMPLEMENTATION

3.1.1 Data Owner

In this module, the data owner uploads their encrypted data in the Cloud server. For the security purpose the data owner encrypts the data file and then store in the server. The Data owner can have capable of manipulating the encrypted data file and performs the following operations Register and Login, Upload File, View Files, Update File, Verify File's Block(Data Integrity Auditing).

3.1.2 Cloud

The Cloud manages which is to provide data storage service for the Data Owners. Data owners encrypt their data files and store them in the Server for sharing with data consumers. To access the shared data files, data consumers download encrypted data files of their interest from the Server and then Server will decrypt them. The server will generate the aggregate key if the end user requests for file authorization to access and performs the following operations such as Login, View and Authorize User, View and Authorize Owner, View Files By Block chain, View All Transactions, Search Requests, Download Requests, View All Attackers, View File Rank Chart, View Time Delay Results, View Throughput Results.

3.1.3 User

In this module, the user can only access the data file with the secret key. The user can search the file for a specified keyword. The data which matches for a particular keyword will be indexed in the cloud server and then response to the end user and performing the following operations Register and Login, Search, Download, View Files, Search Request, Download Request.

3.1.4 TPA

responsible for Login, View File's Meta Data, View Files & Generate Secret Key, CPU Speed.

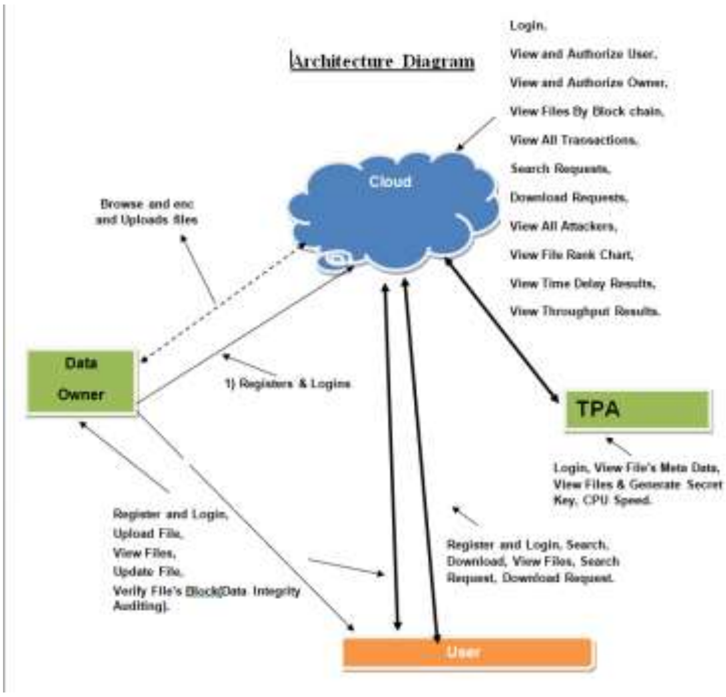
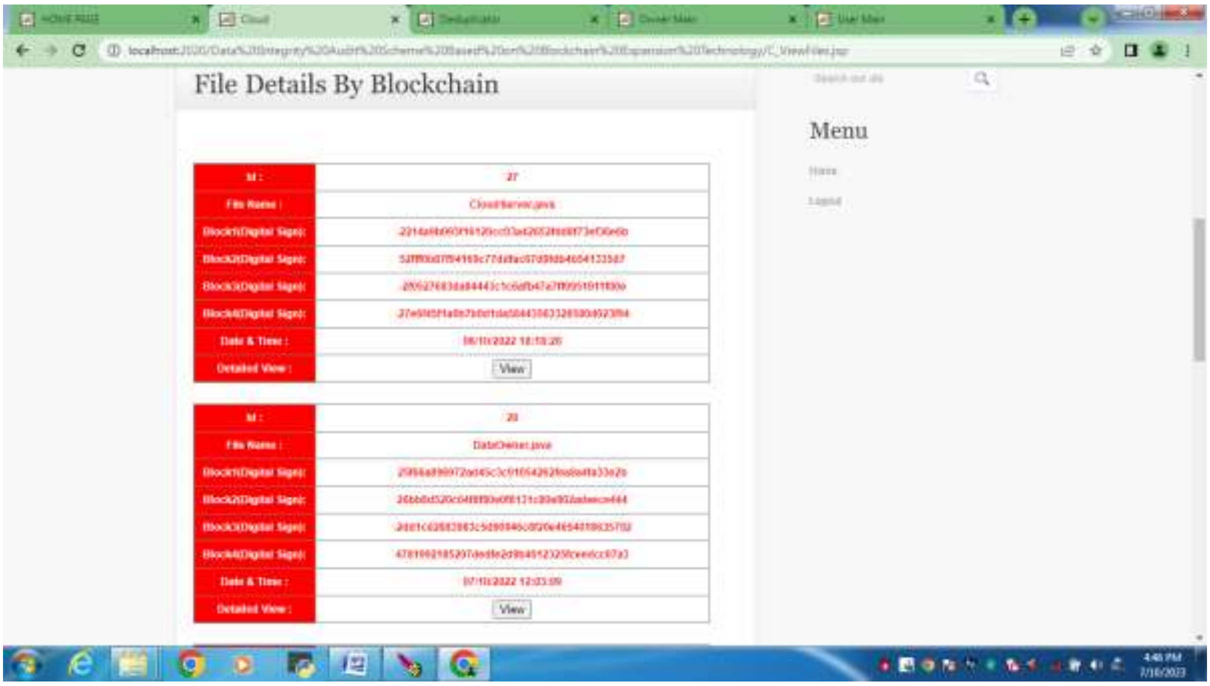


Fig 1: Architecture

4.RESULTS AND DISCUSSION

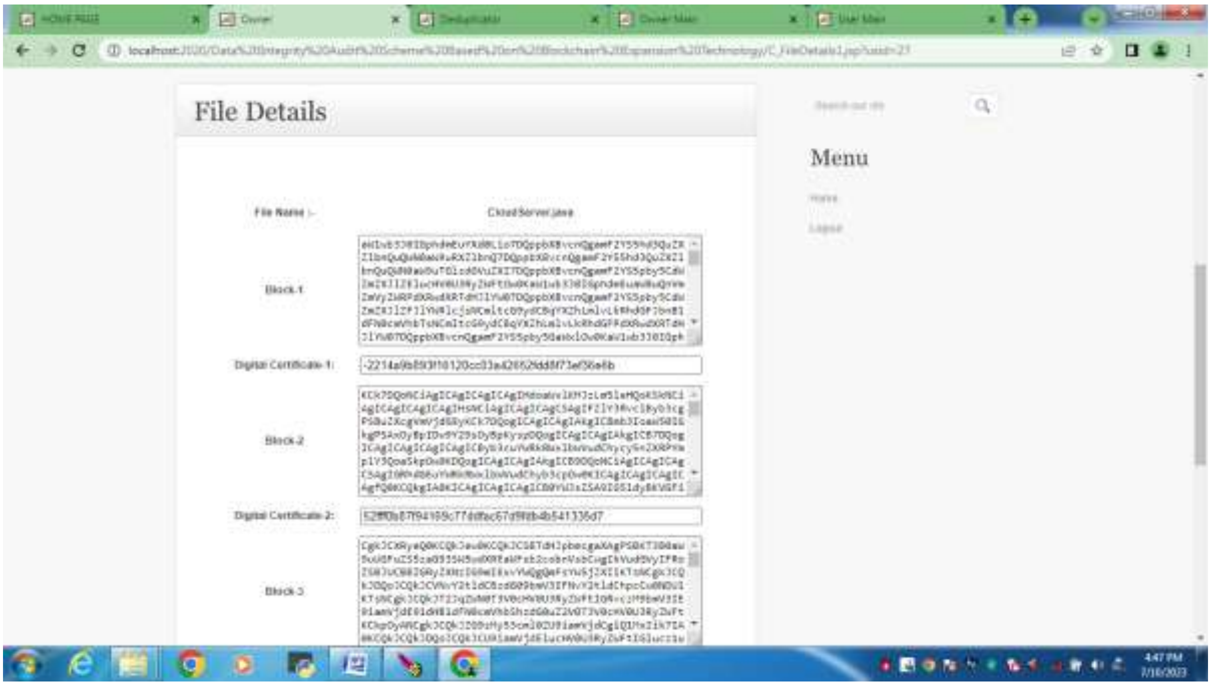


File Details By Blockchain

Id :	27
File Name :	CloudServer.java
Block1(Digital Sign):	-2214a9b893f10120cc03ad2652fdd8f73ef36e6b
Block2(Digital Sign):	52ff0b87f94169c77ddfac67d9fdb4b541336d7
Block3(Digital Sign):	2f0527683da84443c1c6afb47a7ff0951011f00e
Block4(Digital Sign):	27e6f05f1a0b7b0d1da58443063326580d9f023f84
Date & Time :	08/11/2022 18:18:28
Detailed View :	View

Id :	28
File Name :	DataOwner.java
Block1(Digital Sign):	25f66a89972a045c3cf1054262fed6dfa33e2b
Block2(Digital Sign):	26bb8d520c04f8f80a0f8121c80e802adeece444
Block3(Digital Sign):	2d01c62883883c5d90846c8f20e46540f8635702
Block4(Digital Sign):	4781992185297dedfe2d9b4512325fceedcc87a3
Date & Time :	07/11/2022 12:03:09
Detailed View :	View

Fig 2: Uploaded data securely



File Details

File Name :-	CloudServer.java
Block-1	[illegible]
Digital Certificate-1:	-2214a9b893f10120cc03ad2652fdd8f73ef36e6b
Block-2	[illegible]
Digital Certificate-2:	52ff0b87f94169c77ddfac67d9fdb4b541336d7
Block-3	[illegible]

Fig 3: Encrypted data

5.CONCLUSION

As cloud computing and cloud storage technologies improve at an ever-increasing rate and the amount of data saved in cloud storage expands, how can we assure that users have access to all of their data stored on cloud servers? This article suggests a data integrity strategy based on block chain extension technology. We circumvent some of the issues with conventional audits in our plan by using the blockchain network, making it more effective and safe. On the plasma sub-chain as well as the main chain, we deploy smart contracts. The primary chain's capacity strain can be greatly reduced, the development pace can be slowed down, the capacity and computational overhead can be reduced, and the framework execution can be improved thanks to this convention. To ensure the accuracy of the audit and prohibit interaction between the smart contract platform and CSP during contract execution, the reward pool mechanism and the concept of a non-interactive audit are also implemented. The approach can thus accomplish the predicted security objectives.

REFERENCES

[1] K. Hao, J. Xin, Z. Wang, and G. Wang, ``Outsourced data integrity veri_cation based on blockchain in untrusted environment,'' *World Wide Web*, vol. 23, no. 4, pp. 2215_2238, Jul. 2020.

[2] Y. Fan, X. Lin, G. Tan, Y. Zhang, W. Dong, and J. Lei, ``One secure data integrity veri_cation scheme for cloud storage,'' *Future Gener. Comput. Syst.*, vol. 96, pp. 376_385, Jul. 2019.

[3] H. Wang and J. Zhang, ``Blockchain based data integrity veri_cation for large-scale IoT data,'' *IEEE Access*, vol. 7, pp. 164996_165006, 2019.

[4] Z. Miao, C. Ye, P. Yang, R. Liu, B. Liu, and Y. Chen, ``A scheme for electronic evidence sharing based on blockchain and proxy re-encryption,'' in *Proc. 4th Int. Conf. Blockchain Technol. Appl.*, Dec. 2021, pp. 11_16.

[5] F. Kefeng, L. Fei, Y. Haiyang, and Y. Zhen, ``A blockchain-based _exible data auditing scheme for the cloud service,'' *Chin. J. Electron.*, vol. 30, no. 6, pp. 1159_1166, Nov. 2021.

[6] K. He, J. Shi, C. Huang, and X. Hu, ``Blockchain based data integrity veri_cation for cloud storage with T-Merkle tree,'' in *Proc. Int. Conf. Algo-rithms Archit. Parallel Process.* Cham, Switzerland: Springer, Oct. 2020, pp. 65_80.

[7] Y. Lei, Z. Jia, Y. Yang, Y. Cheng, and J. Fu, ``A cloud data access authorization update scheme based on blockchain,'' in *Proc. 3rd Int. Conf. Smart BlockChain (SmartBlock)*, Oct. 2020, pp. 33_38.

[8] Y. Yuan, J. Zhang, W. Xu, and Z. Li, ``Identity-based public data integrity veri_cation scheme in cloud storage system via blockchain,'' *J. Supercomput.*, vol. 78, pp. 8509_8530, Jan. 2022.

[9] S. Wang, D. Zhang, and Y. Zhang, ``Blockchain-based personal health records sharing scheme with data integrity veri_able,'' *IEEE Access*, vol. 7, pp. 102887_102901, 2019.

[10] A. Liu, Y. Wang, and X. Wang, ``Blockchain-based data-driven smart customization,'' in *Data-Driven Engineering Design*. Cham, Switzerland: Springer, 2022, pp. 89_107.

[11] K. Dhyani, J. Mishra, S. Paladhi, and I. S. Thaseen, ``A blockchain-based document veri_cation system for employers,'' in *Proc. Int. Conf. Comput. Intell. Data Eng.* Singapore: Springer, 2022, pp. 123_137.

[12] K. Xu, W. Chen, and Y. Zhang, ``Blockchain-based integrity veri_cation of data migration in multi-cloud storage,'' *J. Phys., Conf. Ser.*, vol.

2132, no. 1, Dec. 2021, Art. no. 012031.

[13] G. Xu, S. Han, Y. Bai, X. Feng, and Y. Gan, ``Data tag replacement algorithm for data integrity veri_cation in cloud storage," *Comput. Secur.*, vol. 103, Apr. 2021, Art. no. 102205.

[14] G. Xie, Y. Liu, G. Xin, and Q. Yang, ``Blockchain-based cloud data integrity veri_cation scheme with high ef_ciency," *Secur. Commun. Netw.*, vol. 2021, pp. 1_15, Apr. 2021.

[15] U. Arjun and S. Vinay, ``Outsourced auditing with data integrity veri_cation scheme (OA-DIV) and dynamic operations for cloud data with multi-copies," *EAI Endorsed Trans. Cloud Syst.*, vol. 7, no. 20, Jul. 2018, Art. no. 169423