International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Issue: 06 | Jun 2023 www.irjet.net p-ISSN: 2395-0072

CLOUD STORAGE AND RETRIEVAL USING BLOCKCHAIN

Mrs. Rashmi M R[1], Raghav Goenka[2], Rohan Sanjeev Tenginkai[3], Vaibhav Singh[4]

[1]*Assistant Professor, Dept. of CSE, NIE, Mysore, Karnataka, India*
[2]*Student, Dept. of CSE, NIE, Mysore, Karnataka, India*
[3]*Student, Dept. of CSE, NIE, Mysore, Karnataka, India*
[4]*Student, Dept. of CSE, NIE, Mysore, Karnataka, India*

Abstract - *Decentralized storage systems leverage distributed architectures and peer-to-peer networks to store data across multiple nodes rather than relying on a single centralized server. This distribution eliminates the vulnerabilities associated with a single point of failure, as well as the risks of data loss or unauthorized access. By dividing data into smaller encrypted chunks and dispersing them across the network, decentralized storage systems ensure data integrity and confidentiality. One of the primary advantages of decentralized storage systems is their enhanced security. With data being encrypted, sharded, and distributed across multiple nodes, it becomes significantly harder for malicious actors to compromise or manipulate the stored information. Furthermore, the absence of a central authority or governing body reduces the risk of data breaches and unauthorized access, empowering individuals, and organizations to have greater control over their data. Decentralized storage systems also offer improved accessibility. Traditional centralized systems often suffer from bottlenecks, limited bandwidth, and geographical restrictions. In contrast, decentralized storage systems allow users to access their data from any node on the network, ensuring faster and more efficient retrieval. Moreover, as these systems rely on distributed networks of volunteers or participants, they can leverage idle storage resources, reducing the costs associated with dedicated infrastructure.*

***Key Words*: Decentralized storage, centralized storage, peer-to-peer networks, security, encryption, distributed network, node etc.**

1.INTRODUCTION

Decentralized storage systems are a type of data storage infrastructure that distributes data across multiple nodes or devices rather than relying on a centralized server or data center. This approach offers several advantages in terms of data security, privacy, availability, and scalability. In a decentralized storage system, data is divided into smaller pieces and encrypted before being distributed across a network of participating nodes. Each node in the network stores a fragment of the data, and multiple copies of each fragment may be stored across different nodes to ensure redundancy and fault tolerance. One key advantage of decentralized storage is enhanced data security. Since data is encrypted and distributed across multiple nodes, it becomes much more difficult for unauthorized parties to access or manipulate the data. Additionally, the use of encryption and fragmentation ensures that even if one node or fragment is compromised, the data remains secure. Decentralized storage systems also offer improved data availability. With traditional centralized storage, if the central server goes down or experiences technical issues, access to the stored data may be lost. In a decentralized system, however, even if some nodes go offline, the data remains accessible as long as enough nodes are still operational. Privacy is another benefit of decentralized storage. Since data is fragmented and encrypted, only the user who possesses the encryption keys can access and decrypt the data. This means that even the storage providers themselves have limited visibility into the actual contents of the stored data. Furthermore, decentralized storage systems are highly scalable. Additional storage capacity can be easily added to the network by simply connecting new nodes, allowing the system to accommodate growing data demands without requiring significant infrastructure upgrades. Some popular examples of decentralized storage systems include InterPlanetary File System (IPFS), Storj, and Filecoin. These systems leverage distributed ledger technologies, such as blockchain, to maintain data integrity, incentivize participation, and facilitate decentralized governance. In summary, decentralized storage systems offer improved data security, privacy, availability, and scalability compared to traditional centralized storage. By distributing data across a network of nodes, these systems provide a robust and resilient approach to storing and accessing data in a decentralized manner.

2.PROBLEM IDENTIFICATION

Data privacy and security are concerns when data resides third party storage. Storage can be created from the underutilized resources of peers. Data security, privacy, availability, and resource utilization are the areas handled by the proposed system.

3.SYSTEM DESIGN AND LIMITATIONS

Zhe, Diao, et al [2], discusses the increasing demand for cloud storage with associated security and privacy issues in centralized cloud storage. As per the discussion by encrypting the data and scattering the data across multiple nodes, a high

level of data security can be achieved. Lee et al [3], has shown encryption enhances the security of user's data stored in cloud storage. Authors have used the AES encryption algorithm to enhance security with speed without impacting the system's performance. Nakamoto, Satoshi [4], uses the concept of bitcoin in blockchain technology to show transaction records. The transaction details are stored in the blocks and are chained to each other serially, using the concept of hashing. Every peer involved in the network has a copy of the blockchain to verify the credibility of the blockchain. The author claims that the transactions stored in the peer-to-peer network are tamperproof, cannot be altered by an attacker and the identity of all the parties involved in transactions is secure. The peerto-peer network uses proof-of-work to record a public history of transactions that quickly becomes computationally impractical for an attacker to change if honest nodes control most of the CPU power. Zyskind [5], raises the problems related to centralized cloud storage and suggest blockchain to solve the issues. The proposed system allows two transactions viz, Taccess and Tdata, Taccess for access control permission which will be set by the respective user who owns the data, and Tdata are used for data storage purposes. The shared encryption key secures the data from third parties. Cachin, Christian et al [6], discusses the architecture of hyper ledger blockchain fabric, limitations of electronic coins, working of hyper ledger fabric, and proof of work consensus algorithm. Hyperledger Fabric is a permissioned blockchain network that allows only limited nodes that have permission to add new blocks in the blockchain. In paper [7], the author explains the architecture of Ethereum and the working of smart contracts. Bitcoin was used only for sending and receiving cryptocurrency but lacked to add business logic. Ethereum applications like decentralized file storage and decentralized autonomous systems are discussed. Ruj, Sushmita [8], proposes BlockStore, a decentralized framework using blockchain technology to enhance security, transparency in transactions between peers (Host and Renters). The system uses proof of storage and proof of work to verify that hosts do not meddle with data in Blockchain. The proposed system does not encrypt or decrypts data before uploading it to peers which creates a threat to confidentiality and privacy of user's data. Juan Benet et al. [9], introduces a new peer to peer file transfer protocol called IPFS (InterPlanetary File System). IPFS uses a content-based addressing scheme. As per the author, IPFS provides a high throughput content-addressed block storage model along with content-addressed hyperlinks. Li, Dagang [10], discusses how data sharing in blockchainbased applications differs from traditional applications. The author identifies that data-sharing in decentralized architecture is cumbersome. The author proposes Meta-key for secure data sharing in a decentralized storage system based on blockchain also focuses on the collusion-free property of the proposed cryptographic protocol and proved it strictly. Wohrer et al. [11], explains the solidity used for creating smart contracts in blockchain and its difficulty. All the security issues which have been resolved are 1. Checks-Effects Interaction, 2. Emergency stop, 3. Speed bump, 4. Rate limit, 5. Mutex and 6. Balance limit. The knowledge related to these issues can be found in grey literature and many blog articles. In [12], blockchain provides scalability, security, and sustainability, it is also helpful to transform the way of doing business. In this paper, the author is trying to conduct a comprehensive survey on the technical and application of blockchain technology by discussing its structure to different consensus algorithms. The author has also explained, the structure of blockchain consists of data, timestamp, and address of the previous block in hash form. The timestamp is recording the time when the block was created. A hash function is the one that takes an input of any length and generates the output with a unique fixed length. Each block contains a hash value of the previous block. therefore, security is increased in Blockchain. It uses proof of stack (POS), proof of work (POW) consensus algorithm as a measure to discourage the attacks of Denial of Service and miner can validate transactions in a block depending on the amount the user holds respectively. Therefore, blockchain technology is exceeding recognized and appraised due to its decentralized infrastructure and peer-to-peer nature. In paper [13] D. Sivaganesan has suggested the use of blockchain to improve security and provide transparency in IoT applications. The use of smart contract avoids third person intervention as well as provides improved security and transparency in the transaction.

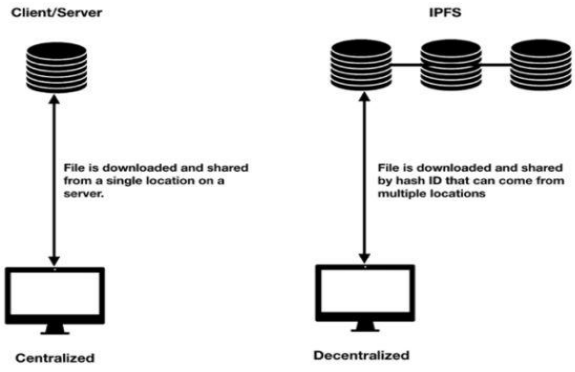

Fig 1. Difference between Centralized and Decentralized

In the proposed system, smart contracts are used to store file details in the blockchain and transfer the cryptocurrency from the user's wallet to the peer's wallet. AES encryption algorithm for enhancing the security of user's data stored in cloud storage. The proposed system maps the user's wallet address with the user's file so that only the legitimate owner can access the file's data. Users' information is stored in Ethereum blockchain. The Ethereum blockchain network allows the use of smart contracts through which information of file uploaded by the user is stored in the blockchain. The proposed system encrypts and decrypts data every single time for upload and download operation. The system uses the IPFS Protocol to distribute files efficiently across several peers in the network.

4. METHODOLOGY

The proposed system works in four modules. The user first creates an account on the metamask. The user's account address and wallet balance are fetched in the app through web3.js from the metamask. Users select the file to upload through file picker. System checks for the number of available peers. Further, the AES algorithm uses the user's wallet address as a key and encrypts the uploaded file. A payment dialogue seeks for the user's confirmation. On confirming the payment, the user's file is stored across available peers using IPFS protocol. IPFS then returns a hash value consisting of the path of the file. This path is then mapped with the user's address using a smart contract and gets stored securely in the blockchain. To achieve high availability and reliability of data, the uploaded data is replicated on three peers. For better performance system blacklists peers every time they are unavailable for data retrieval.

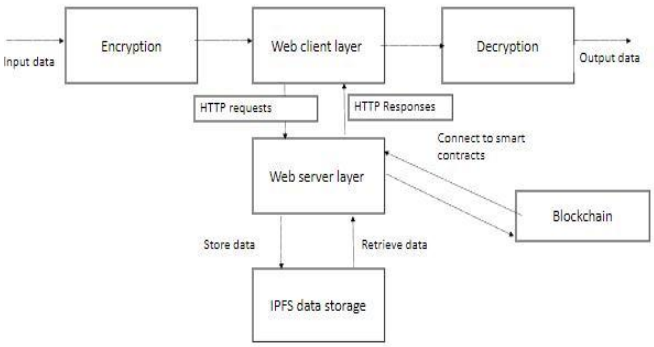

Fig 2. System Design for Cloud Storage & Retrieval

The terminology is briefly discussed below:

1. Metamask: Browser extension which acts as a bridge to connect with the ethereum network.

2. Ethereum network: It is an open-source, public blockchain based distributed computing platform. Ethereum uses smart contracts where one can add business logic to make decentralized applications as per the business requirements.
3. Peers: These are the users of the system who have pledged to rent their free storage for other users to store files.

4. AES: Advance Encryption Standard (AES) is a symmetric key algorithm that supports block length of 128 bit and can have a key size of 128, 192, and 256 bits.

5. IPFS protocol: IPFS is an open-source peer to peer file transfer protocol.

A. Uploading of file: User uploads file using the file picker. The system checks the file size and ensures storage availability in the network. The file is uploaded when enough storage is available. Then system performs step B. Users are notified to try again when enough storage is unavailable.

B. Encryption of file: The uploaded file is encrypted using AES 256-bit algorithm. The encryption key is generated using the user's wallet address and randomly generated salt value. This encryption key along with an IV is used to encrypt user's data. This maintains the confidentiality of the user's data.

C. Storing of file across multiple peers: The encrypted file is then divided into blocks of 64KB and sends to different peers across the network with the help of the IPFS protocol. The proposed system uses a private IPFS network to allow registered peers to store the file in the network. The file block is replicated on multiple peer's storages for high availability using the IPFS cluster.

D. Storing of file across multiple peers: IPFS returns a hash value which indicates the path of the file. The hash value along with metadata is mapped with the user's wallet address and is stored in the blockchain using a smart contract. Smart contracts are like agreement and are used to eradicate the need for a third party. They control the transaction between nodes or assets between parties under certain conditions. This is lines of code stored on a blockchain network and are automatically executed when predetermined terms and conditions are met. In our proposed system preconditions for the smart contract to execute are: 1) Enough Space is available in the network to store files. 2) The user has sufficient wallet balance to pay the peers.

The smart contract stores all the files details in the structure named File and maps this structure with the user's address. It consists of two functions, one to add a new file and another to get the details of the uploaded file.

E. Paying the peers for file storage: Once the file is stored across peers, total cryptocurrency is calculated and is deducted from the user's wallet. This cryptocurrency is first transmitted to the smart contract from the user's wallet. With the smart contract, this amount is distributed to the peers who have stored the user's file.

Fig 3. Smart contract to store file details

The above smart contract stores all the files details in the structure named File and maps this structure with the user's address. It consists of two functions, one to add a new file and another to get the details of the uploaded file.

5. RESULTS

To access the system, users first sign up on metamask and login with the registered credentials. Successful login takes users to the home screen for selecting the file to upload.

Fig 4. GUI to upload the files

System checks for storage availability based on selected file size. The selected file is encrypted using AES 256-bit algorithm when sufficient storage is available. The system will compute the total cost of storing the file. Once the cost is the calculated system will check if the user's wallet balance is more than the calculated cost. If the user has sufficient balance, then he/she is prompt to pay the cryptocurrency to store the file.

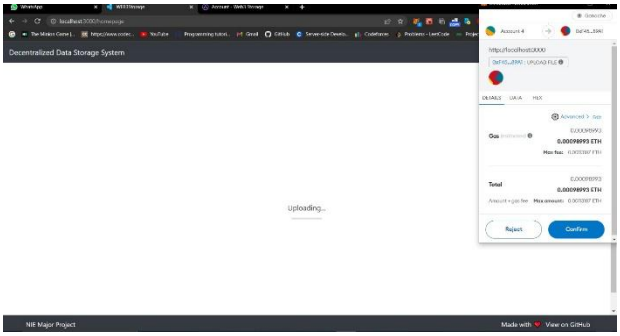

Fig 5. Payment Confirmation using Metamask

After a successful payment file is split into blocks and store across peers using IPFS protocol and the corresponding hash value is stored in the blockchain.

Fig. 6 represents successfully uploaded file details. Once the file is successfully uploaded ipfs returns a hash value indicating the path of the file. This will be mapped with the user's wallet address and will be stored in the blockchain with the help of a smart contract.

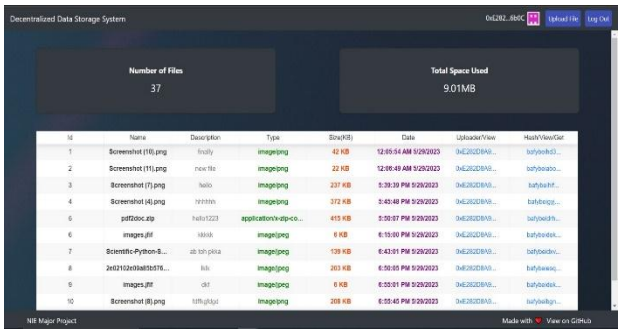

Fig 6. File uploaded successfully

6. CONCLUSIONS

The proposed system enhances the security of data by encrypting and distributing the data across multiple peers in the system. Implemented system uses the AES 256bit encryption algorithm to encrypt the data ensuring the confidentiality of the user's data. Encrypted data is then distributed and stored across peers in the network using the IPFS protocol. Our system not only solves the privacy and security concerns of centralized cloud storage but also provides a medium for the peer to rent their underutilized storage and earn cryptocurrency in return thereby, maximizing the storage resource utilization. In the future, an adaptive scheduling algorithm can be incorporated with which files can be accessed multiple times by the user as compared to the one which is accessed rarely. This will help to ensure that frequently accessed files are available easily to the user whenever required. Also, a credit system can be added with which each peer will be assigned a default 100 credit, based on their system uptime, and several successfully served file access that requests their credits will

be either deducted or added. Peers with more credits will be given higher priority for data storage.

REFERENCES

[1] Bernard Marr," How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read." Forbes, 2018 M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[2] Zhe, Diao," Study on Data Security Policy Based On Cloud Storage" 2017 ieee 3rd international conference on big data security on cloud (bigdatasecurity), ieee international conference on high performance and smart computing (hpsc), and ieee international conference on intelligent data and security (ids) IEEE, 2017.

[3] Lee, Bih-Hwang, Ervin Kusuma Dewi, and Muhammad Farid Wajdi. " Data security in cloud computing using AES under HEROKU cloud." 2018 27th Wireless and Optical Communication Conference (WOCC).IEEE, 2018.

[4] Nakamoto, Satoshi,"Bitcoin: A peer-to-peer electronic cash system", (2008).

[5] Zyskind, Guy, and Oz Nathan," privacy: Using blockchain to protect personal data",IEEE Security and Privacy Workshops. IEEE, 2015

[6] Cachin, Christian," Architecture of the hyperledger blockchain fabric", Workshop on distributed cryptocurrencies and consensus ledgers. Vol. 310. 2016.

[7] Buterin, Vitalik," A next-generation smart contract and decentralized application platform", white paper (2014).

[8] Ruj, Sushmita, et al," BlockStore: A Secure Decentralized Storage Framework on Blockchain" 2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA). IEEE, 2018.

[9] Benet, Juan, "IPFS - Content Addressed, Versioned, P2P File System." 2014.

[10] Li, Dagang, et al. Meta-Key:" A Secure Data-Sharing Protocol Under Blockchain-Based Decentralized Storage Architecture", IEEE Networking Letters 1.1 (2019): 30-33.

[11] Wohrer, Maximilian, and Uwe Zdun,"a Smart contracts: security patterns in the ethereum ecosystem and solidity", International Workshop on Blockchain Oriented Software Engineering (IWBOSE). IEEE, 2018.

[12] Sum, V. "SECURITY AND PRIVACY MECHANISM USING BLOCKCHAIN." Journal of Ubiquitous Computing and Communication Technologies (UCCT) 1.01 (2019): 45-54

[13] Sivaganesan, D. "BLOCK CHAIN ENABLED INTERNET OF THINGS." Journal of Information Technology 1.01 (2019): 1-8.