Quest Journals
Journal of Software Engineering and Simulation
Volume 10 ~ Issue 7 (2024) pp: 27-31
ISSN(Online) :2321-3795 ISSN (Print):2321-3809
www.questjournals.org

Research Paper

Comprehensive Review of Decentralized Cloud Storage Using Blockchain

Malavika N
Dept.Computer Science and EngineeringWYD23CSNS05

Hasna M
Dept.Computer Science and Engineering

Abstract*—The cloud storage system has become the one of the most important aspect in everyone's lives.It provide a convenient and efficient way to store and access the data to all the users.As the cloud storage is centralized it may have many drawbacks such as privacy concern, data security and lack of control over the stored data. These issues can be solved using a decentralized cloud storage using blockchain. This system will provide the users with secure, transparent and user controlled storage of data. It also provide the users to utilize the remaining space of hard disk.Any of the computing node which is connected to the internet can join and can form peer network by maximizing the resource utilization.*

Keywords: *Decentralized; encryption; decryption;blockchain; cloud computing*

***R**eceived 01 July, 2024; **R**evised 09 July, 2024; **A**ccepted 11 July, 2024 © The author(s) 2024.*
P**ublished with open access at **www.questjournals.org

I. INTRODUCTION

The centralized cloud storage offers convenience and cost effectiveness, it comes with security and transparency con- cerns. Decentralized cloud storage provides improved security,transparency, and user control but may face challenges in terms of complexity and scalability. Centralized cloud storage systems are more vulnerable to hacking, data breaches, and unauthorized access, that results in the loss of sensitive data. Decentralized cloud storage systems can solve these issues by distributing the data across a network of nodes, so by makingit more secure, transparent, and user-controlled. The system will provide high level of security for users the stored data. Blockchain provides transparency and immutability which ensures the stored data is secured and it cannot be altered without the permission. The system also ensures privacy tothe users stored data. By encrypting the data using the user's public key, the user can access and decrypt the data using their private key. Blockchain based cloud storage system can audit the storage and retrieval of the data. This approach will provide transparency to the users. A Decentralized Cloud Storage ap- plication can offer a range of advantages, addressing key issues related to data security, privacy, cost, and reliability. As the importance of data continues to grow, solutions that prioritize these aspects become increasingly valuable for individuals and organizations alike. Some of the applications provided by decentralized cloud storage system are enhanced security, cost effectiveness, improved reliability, increased privacy, data ownership and control,scalability, censorship resistance, com- munity participation and reduction of environmental impact.

II. LITERATURE REVIEW

A. Decentralized Cloud Storage Using Blockchain

[8] The proposed system operates through four distinct modules, as depicted in the accompanying figure. Users begin by creating an account on MetaMask, which allows the appto retrieve the user's account address and wallet balance using web3.js. Upon selecting a file for upload via a filepicker, the system checks the availability of peers. The AES algorithm then encrypts the uploaded file using the user's wallet address as the key. A payment dialogue prompts the userfor confirmation; once confirmed, the file is securely stored across available peers using the IPFS protocol. The IPFS returns a hash value that includes the file path, which is then linked to the user's address through a smart contract, ensuring secure storage on the blockchain. To enhance data availability and reliability, the system replicates the uploaded data across three separate peers. Additionally, to optimize performance, the system blacklists peers each time they become unavailable for data

retrieval.

B. Blockchain-Based Decentralized Cloud Solutions for Data Transfer

[9] This study focuses on addressing the challenges as- sociated with transferring data, along with the relevant per- mission poli- cies, from the cloud to the Distributed File System (DFS). In the DFS, no specific node is privileged, and data storage relies on content addressing. Ensuring protection against unauthorized access, the system implements role-basedor user-based access control, akin to major cloud providers' offerings. To elaborate on the proposed strategy's components,a mapping is conducted with the current DFS and cloud provider. AWS is selected as the cloud provider, while the InterPlanetary File System (IPFS) serves as the DFS. Thedata files transition from Amazon S3 buckets to IPFS, and the authorization policies associated with S3 buckets, specifically the resource-based IAM rules, are leveraged for implementingaccess control on IPFS files. This technique is designed tobe adaptable for use with other cloud providers and DFS. Amazon Simple Storage Service (S3) and Amazon Identity and Access Management (IAM) are employed. AWS IAM manages permission and authentication components, allowing administrators to configure users and groups with access or restrictions to resources. Amazon S3 provides scalable storage,with data organized into buckets, each subject to access controls. The data files are then moved from S3 buckets to IPFS, distributed across the IPFS network for improved scalability, availability, and security.

C. Blockchain-based Decentralized Storage Scheme

[7] Blockchain employs a distinctive data structure that con- nects data blocks in a sequential, chronological order. This structure ensures decentralization, immutability, and a dis-tributed ledger system that cannot be tampered with or forged.It leverages a distributed system architecture incorporatingtechniques like hashing, asymmetric encryption, Merkle trees, and timestamps.

Lightning Network: The Lightning Network primarily uses two types of contracts: Sequence Expiration Revocable (RSMC) and Hash Time Locked Contract (HTLC). These con-tracts address the issues of rapid two-way and node-to-node transfers. The Lightning Network supports one-way payments between two users and leverages the sequence expiration revocable contract for unlimited fast offline transfers.

Data Integrity: In the blockchain-based distributed storage system, a Merkle tree-based data integrity verification scheme is adopted. Users encrypt their file data and generate random challenges corresponding to each data block. These challenges,along with the block data, are hashed to create Merkle leaf nodes, forming a Merkle tree. During the verification phase, users randomly select challenges and send them to intermedi- aries to validate the data's integrity.

D. Block Chain Based Decentralized Cloud Storage

[8] As decentralized storage models become more integral to blockchain technology, there is an increasing trend toward using a blockchain to store the root of the Merkle tree. The Merkle tree is an efficient way of organizing large data setsby combining data into blocks, hashing these blocks, and then iteratively hashing the results until a single root hash is produced. This root hash is subsequently stored on the blockchain. The Merkle tree is structured as a binary tree, where each node consistently contains two blocks or hashes. By utilizing a blockchain that maintains the root of numerous data elements, it becomes possible to verify the existence of millions of files within a single transaction.

The encryption method employed is authenticated encryp- tion, supporting both the AES-GCM cipher and the combina- tion of Salsa20 and Poly1305, known as "Secretbox" in NaCl. Authenticated encryption allows users to detect any tamperingwith their data.

Key aspects of the encryption process include:

* Data encryption is performed in small batches, typically recommended to be 4KB or less.
* While the same encryption key is used for each encryptionbatch within a segment, different segments may use different encryption keys.
* The nonce for each encryption batch must increment monotonically from the previous batch throughout the entire segment. If the counter reaches the maximum representable nonce, it wraps around to 0.
* To prevent reordering attacks, the starting nonce of each segment is deterministically chosen based on the segment number.
* In scenarios involving the simultaneous upload of multiple segments, such as with Amazon S3's multipart-upload feature,the starting nonce for each segment can be calculated from thefile's starting nonce and the segment number.
* This encryption scheme ensures that the content of the data is protected from the storage node that holds the data. The data owner retains full control over the encryption key, thus maintaining authority over data access.

By leveraging these techniques, the system ensures ro- bust security and integrity for decentralized storage within a blockchain framework.

E. Decentralized Cloud Storage System Using Blockchain

[6] The system design for the decentralized cloud stor- age system, incorporating blockchain, Hardhat, and Pinata, involves the integration of various components to establish a secure and efficient platform for storing and retrieving data files. The subsequent sections elaborate on the architecture, smart contracts, blockchain network, storage nodes, front-end application, Hardhat, and Pinata. The architecture of the decen-tralized cloud storage system comprises a blockchain network, a front-end application, and storage nodes. The blockchain network functions as the system's backbone, storing file meta- data and executing smart contracts. The front-end application provides a user interface for file upload and retrieval, while storage nodes store the actual data files. Smart contracts for the decentralized cloud storage system include a file registry contract and a storage contract. The file registry contract stores metadata such as file name, size, and hash, while the storage contract manages storage nodes and their capacity. Developed in Solidity, these smart contracts operate on the Ethereum blockchain. The Ethereum network serves as the blockchain network for the decentralized cloud storage system. Renowned for its decentralization, security, and immutability, Ethereum is suitable for storing file metadata and executing smart contracts. The scalability of the Ethereum network allows the system to handle a large volume of transactions. Storage nodes are responsible for storing actual data files.

Decentralized and connected to the blockchain network, these nodes communicate with smart contracts for data retrieval and storage. Designed to be decentralized, no single node controls the data files, and these nodes can be various de- vices with internet connectivity and sufficient storage capacity Decentralized cloud storage is designed around blockchain, with IPFS (Interplanetary File System) used to store data on the network. Pinata services assist in storing data on IPFS. Ethereum is employed as the blockchain for storing smart contracts. Hardhat, an open-source development environment built on Node.js, aids in building and testing smart con- tracts for Ethereum-based decentralized applications (dApps). Data collection encompassed user engagement, performance metrics, security, blockchain metrics, and user demographics. Statistical analysis and data visualization techniques were applied to optimize performance, enhance user experience, and develop targeted marketing campaigns. For instance, user en- gagement data identified popular features, while performance metrics pinpointed bottlenecks for optimization. Security data ensured data protection, and blockchain metrics optimized blockchain performance. User demographics guided targeted marketing efforts. The results informed strategic improvements and optimizations across the application.

F. Decentralized Cloud Storage Using Blockchain

[10] The process involves data owners registering them- selves, logging in, and uploading files. The system ensures file size compliance and checks storage availability before ini-tiating the upload process. Subsequently, the system performs several key steps:

- Encryption: The uploaded file undergoes encryption using the AES 256-bit algorithm. The encryption key is generated using the owner's wallet address and a randomly generated salt value, ensuring data confidentiality.
- File Division and Distribution: The encrypted file is divided into blocks of 64KB, and these blocks are sent to different peers across the network using the IPFS protocol. A private IPFS network is utilized, allowing registered peers to store files within the network.
- Hash Mapping and Blockchain Storage: The system gener-ates a hash value representing the file's path. This hash value, along with metadata, is mapped to the user's wallet address and stored in the blockchain using a smart contract. Smart contracts eliminate the need for third-party intermediaries, automating agreements under specific conditions.
- Access Control: Only the user possessing the hash values stored in the blockchain can combine the small files into one large file. Users must register and log in, and access is granted only if the file owner permits it. Users can search for files and request access, with access granted by the owner.
- Data Traceability: The cloud storage system maintains information about the data owner, data user, block data, and can trace data. Hash values from previous and current states allow tracking of the peer where the data is stored.

The internal operations involve interactions with the front end, which are submitted to a controller for information validation and forwarding, resulting in data changes. Con-sensus algorithms run and examine each block, with the P2P network's listings updated in case of discrepancies. Otherwise, the block is accepted, and adjustments are not made. The transaction list is updated, and the new block is inserted using the IPFS protocol. The project aims to create a secure and efficient decentralized storage system, leveraging blockchain technology

and related protocols.

G. A Blockchain-based Decentralized Data Storage andAccess Framework for PingER [1]

In the PingER Monitoring Agent (MA) system, each setof sample measurements is sent every 30 minutes. The MA performs its monitoring task by sequentially sending up to thirty 100-byte pings at one-second intervals to remote sites. This continues until it receives ten echo replies or times out after 30 seconds. The same process is repeated for 1000-byte pings. The collected data for each ping set includes details such as the MA name, remote site information, IPaddresses, payload in ping requests, timestamp, the numberof ping packets sent and received, response sequence number, and Round Trip Time (RTT) statistics (minimum, average, and maximum values). All these raw measurements are stored in flat text files at each MA.

A centralized data repository at SLAC collects these text archives from each MA daily. The retrieved data is then analyzed and stored using a specific naming convention that includes the performance metric name, packet size, MA name,and the measurement date.

The processed data, referred to as hourly data from all MAs, is used to generate sixteen Internet performance metrics on a daily, monthly, and yearly basis. This data is publicly accessi- ble and can be downloaded from the pingtable web interface on the SLAC web server or via anonymous FTP. Thus, the current PingER framework relies on centralized storage and processing, with SLAC resources handling analysis, archiving,and reporting.

H. Research on Decentralized Storage Based on a Blockchain

[3] The decentralized storage framework presented includes four primary components: the user layer, data processing layer,storage network layer, and blockchain layer. The user layer supports user registration and differentiates between regular users and administrators. The system incorporates an access system where ordinary user registrations require administrator approval. The data processing layer handles various tasks, including data uploading, downloading, encryption, and frag- mentation.

In the storage network layer, the IPFS protocol is employed, leveraging P2P technology to connect nodes within the net- work and store the processed data fragments. The blockchain layer uses a federated chain and interacts with the external environment via smart contracts. This layer manages metadatauploads, including file hash, file name, owner information, and more. Users can access the stored file metadata on the blockchain to retrieve their files.

During data storage, users encrypt their data files usinga combination of the RSA algorithm and AES double en- cryption. The encrypted files are then fragmented according to specific rules and uploaded to the IPFS network. The IPFS network generates a hash for the uploaded file, and this information, along with the file name and owner details, is stored as a JSON file on the blockchain.

For data retrieval, users first obtain the file's storage address.The IPFS network's storage address is determined by the file's contents, requiring only the file hash for retrieval. Users verify the hash value of the data on the blockchain, retrieve theencrypted data from the IPFS network using the hash value, and finally decrypt the data to access the original content.

I. Blockchain-Based Decentralized Cloud Storage

[6] In this paper, a public audit scheme utilizing blockchain technology to address issues related to malicious auditors [6]. However, the introduced plan faces a significant challenge where customers require access to complete data backups,which is impractical in real-world scenarios. In practical appli-cations, the task of integrity checking is typically undertaken by a Third-Party Auditor (TPA), and many later-proposedschemes support public auditing. Upon analysis, it identified certain drawbacks in the proposed scheme, particularly its dependence on a trusted third party, the TPA. To validatethe effectiveness of approach, a prototype on the Ethereum platform, leveraging Aliyun as a data storage service is developed. Performance tests were conducted for uploading and downloading files of various sizes. It's worth noting that blockchain imposes limitations on block capacity, necessitat- ing the storage of only highly sensitive security information in blocks to prevent a significant impact on system performance. While analyzing user operations from extensive system logs is deemed inefficient, the metadata information on the blockchaincan be utilized later for verifying data authenticity and trackingfile access. Auditing operations can be efficiently conducted through the analysis of the operations log. The implementation phase is a crucial step in the system life cycle, where the theoretical design is transformed into a functional system. This final phase ensures the conversion of the new system into an operational one. The testing process consists of two key components: Unit testing involves a set of tests performed by individual programmers before integrating the unit into a larger system. The module interface is tested to ensure proper information flow in and out of the program unit. Local data structure examination ensures the temporary storage integrity during algorithm execution. Boundary conditions are tested to ensure the proper operation of the module within established limits. All independent paths through the control

structure are tested, including error-handling paths. Black-box testing is a comprehensive method examining the functionality of an application without delving into its internal structures. This method is applicable to various levels of software testing,including unit, integration, system, and acceptance testing.

Also known as specification-based testing, it focuses on vali- dating the software against specified requirements. The systemimplementation phase, therefore, marks the transition fromtheoretical design to a fully operational system, ensuring the practicality and effectiveness of the proposed blockchain-based public audit scheme.

III. CONCLUSION

Cloud computing systems are widely adopted for data sharing across numerous applications and network compo-nents. However, the decentralized nature of cloud systems, with multiple copies of data taking various paths to ensure resilience, poses challenges for administrators in identifying the origin, impact, and tools involved in potential security breaches. The fundamental pillars of blockchain architecture, combining cryptographic mechanisms with distributed public ledgers, offer a promising solution to address these challenges. This amalgamation facilitates the creation of diverse structures on a blockchain without introducing trust issues within the net- work. The application of blockchain in cloud solutions, partic- ularly using the Ethereum platform, brings several advantages, including data provenance verification (ensuring the data's source) and enhanced cloud monitoring. In a blockchain- enabled cloud environment, the presence of genuine data provenance across cloud servers, distributed data calculations, transfers, and transactions serves as a robust mechanism for detecting insider threats, validating test findings, and pinpoint- ing the specific source of a system or network breach. The use of blockchain has become a crucial technique for ensuring se- curity, emphasizing integrity, authenticity, and confidentiality. The identified benefits of blockchain in cloud computing have motivated us to explore and contribute to security measures in this context, fostering a more resilient and trustworthy cloud infrastructure.

REFERENCES

[1]. Yan Zhu et al, Yan Zhu1, Chunli Lv1, Zichuan Zeng1, Jingfu Wang1 , Bei Pei2. "Blockchain-based Decentralized Storage Scheme" 2019 J.Phys.: Conf. Ser. 1237 042008
[2]. N.M.K. Ramalingamsakthivelan, B. Ramya. "Blockchain-Based Decen-tralized Cloud Storage" 2023 Volume 10, Issue 4
[3]. Lu Meng, Bin Sun. "Research on Decentralized Storage Based on a Blockchain" 2022
[4]. Nabeel Khan, Hanan Aljoaey, Mujahid Tabassum, Ali Farzamnia, Tripti Sharma , Yew Hoe Tung."Proposed Model for Secured Data Storage in Decentralized Cloud by Blockchain Ethereum " 2022
[5]. Saqib Ali†, Guojun Wang , Bebo White, Roger Leslie Cottrell "A Blockchain-based Decentralized Data Storage and Access Framework for PingER" 2018
[6]. G. Richa Shalom, Ganesh Rohit Nirogi."Decentralized Cloud Storage Using Blockchain" Volume 10 Issue IX Sep 2022
[7]. G. Abinaya, Preksha Kothari, Alex Pavithran KP, Manasi Biswas, Farheen Khan "Block Chain Based Decentralized Cloud Storage" , Volume-8 Issue-4, April 2019
[8]. Jeppiaar Nagar, Rajiv Gandhi Salai."DECENTRALIZED CLOUD STORAGE USING BLOCKCHAIN" 2021
[9]. Rajit Nair , Syed Nasrullah Zafrullah , P. Vinayasree , Prabhdeep Singh ,Musaddak Maher Abdul Zahra , Tripti Sharma , and FardinAhmadi."Blockchain-Based Decentralized Cloud Solutions for Data Transfer"2022
[10]. Meet Shah ,Mohammedhasan Shaikh ,Vishwajeet Mishra,Grinal Tus- cano."Decentralized Cloud Storage Using Blockchain" 2020