Decentralized File System (Storage and Sharing) Using Blockchain

Anurag Rathour[1], Aditya Shahi[2], Ashutosh Tiwari[3], Babulal Maurya[4], Manish Jha[5]

[1] Student, Information Technology, Institute of Technology and Management Gorakhpur, Utter Pradesh, India
[2] Student, Information Technology, Institute of Technology and Management Gorakhpur, Utter Pradesh, India
[3] Student, Information Technology, Institute of Technology and Management Gorakhpur, Utter Pradesh, India
[4] Student, Information Technology, Institute of Technology and Management Gorakhpur, Utter Pradesh, India
[5] Professor, Information Technology, Institute of Technology and Management Gorakhpur, Utter Pradesh, India

ABSTRACT

The introduction starts by pointing out the increasing importance of file sharing and storage in the digital age, with an emphasis on the traditional reliance on centralized systems. Discuss the vulnerabilities and limitations of these systems, such as a single point of failure and susceptibility to security breaches. Then, introduce the concept of decentralization and the role blockchain plays in transforming file sharing and storage. Emphasize the potential of blockchain to address the shortcomings of centralized systems by providing enhanced security, transparency, and trust. This paper focuses on decentralized secure data storage and sharing, high availability of data, and efficient utilization of storage and sharing resources.

Keyword- *Blockchain*, *IPFS, Pinata, MetaMask etc.*

INTRODUCTION-

In the realm of digital information management, traditional centralized cloud storage systems face escalating concerns over security vulnerabilities and a susceptibility to breaches. While cloud storage remains a primary option for handling extensive datasets, the drawbacks of centralized control have prompted a paradigm shift. Blockchain technology has emerged as a transformative force, introducing decentralized file sharing and storage systems that fundamentally alter the landscape of data management. The essence of this shift lies in moving away from centralized models, where data resides in single points of control, towards a decentralized framework facilitated by blockchain. Inherently securing data, blockchain operates as a distributed ledger technology, establishing a decentralized cloud storage system. This innovation hinges on the formation of a peer-to peer network, allowing any connected computing node to actively engage and optimize resource utilization. A pivotal player in this transformation is the Interplanetary File System (IPFS), a protocol that enhances the security and efficiency of file storage on multiple network peers. In this proposed system, user files undergo encryption and distribution across the IPFS network, with hash values pointing to file paths stored on the blockchain. This decentralized and encrypted approach not only addresses security concerns but also empowers users with greater control and privacy over their data. Nevertheless, the integration of large files into the blockchain ecosystem presents distinctive challenges. Inherent limitations such as block size constraints and the necessity to split and reassemble files off-chain introduce inefficiencies. Blockchain bloat further exacerbates the situation by replicating data across numerous nodes, resulting in heightened storage demands and operational costs. To mitigate these challenges, a proposed modification to IPFS incorporates Ethereum smart contracts, 1 2 enabling access-controlled file sharing. Smart contracts facilitate the maintenance of access control lists, ensuring that file-sharing activities align with predefined security parameters. The fusion of blockchain and decentralized file sharing not only addresses security concerns but also introduces heightened transparency, trust, and user autonomy. By leveraging cryptographic hash functions within IPFS and capitalizing on the decentralized nature

of blockchain, users gain the ability to securely share files within a tamper-proof and thrustless environment. This review paper undertakes a comprehensive exploration of the intricate relationship between blockchain technology and decentralized file sharing within a concise framework. It examines strategies, advantages, and challenges associated with this transformative technology, shedding light on its potential applications and implications for the future of secure and efficient data storage. Throughout the paper, insights into existing decentralized file-sharing platforms will be provided, offering a nuanced understanding of their functionalities and real-world applications.

PROBLEM STATEMENT

Traditional file-sharing and storage systems have inherent weaknesses due to their centralized structure, making them susceptible to security breaches, limitations, and system failures. When a centralized system experiences a breakdown, it risks the loss of vital data, and the hefty maintenance costs often result in expensive service charges for users. However, decentralized file sharing and storage systems, operating on Blockchain technology, aim to tackle these challenges by offering a secure, dependable, and transparent platform resilient to failures and malicious attacks. Through Blockchain-based decentralized systems, peer-to-peer interactions facilitate seamless sharing of information, alleviating concerns about potential data loss or system downtime. Moreover, Blockchain technology ensures data security by making each transaction transparent, tamper-proof, and immutable, thereby enhancing trust and reliability in the system.

By decentralizing file sharing and storage processes, Blockchain technology empowers users with greater control over their data while significantly reducing the risks associated with centralized systems. This innovative approach not only enhances security and reliability but also fosters a more efficient and transparent sharing environment. As such, Blockchain-based decentralized file sharing and storage systems represent a paradigm shift in the way data is managed and exchanged, offering promising solutions to the challenges faced by traditional centralized systems.

METHODOLOGY

In this system, users are given access using public keys for transactions, ensuring only authorized individuals can access data through public key cryptography. Files are stored using Piñata, an IPFS platform, providing decentralized storage accessible from anywhere. Development involves creating a user-friendly interface, integrating with IPFS, and writing smart contracts for smooth operation. Tools like React.js and ether.js are used for the frontend, while Solidity is employed for smart contract writing, enhancing security and efficiency.

Transactions are authenticated using MetaMask wallets, ensuring only authorized users can complete them. Hardhat simplifies smart contract deployment and testing, reducing development time and improving reliability. Overall, this process ensures the file sharing system is secure, reliable, and easy to use for authorized users

LITERATURE REVIEW-

[1] DECENTRALIZED CLOUD STORAGE USING BLOCKCHAIN-

Author - G. Richa Shalom, Ganesh Rohit Nirogi.

This article describes a way to use blockchain applications to interact directly with blockchain or smart contracts to obtain transaction confirmation, information, or execution. A heterogeneous network of nodes stores the blockchain, processes transactions, and executes smart contracts when necessary. This may cause the following problems when processing large files. Since this information generally does not need to be run on the blockchain node, the blockchain becomes bloated, causing information to be copied to multiple nodes for analysis. A blockchain has a series of transactions logically arranged in a block that are linked together using cryptography

to form a chain. The technology itself creates a common consensus that allows peers to reach a consensus on the status of the data transfer exchange. The technology itself creates a common consensus that allows peers to reach a consensus on the status of the data transfer exchange. The technology itself creates a common consensus that allows peers to reach a consensus on the status of the data transfer exchange. Information is usually stored in special files maintained by each organization. Information technology is often defined according to the CAP theorem, including consistency (C), availability (A), and distributed tolerance (P). No database, whether SQL or non-SQL, can simultaneously implement all three properties. In the case of blockchain, data is organized using linked lists using hash pointers instead of plain pointers. The transaction has size variables and includes a reference to the previous block if it is not a genesis block. The genesis block is the first block of the blockchain that was hardcoded at the time of the release of the blockchain, and its structure varies depending on which blockchain technology is used requires an address to add a block to the blockchain. The address is derived from public key. These are unique identifiers that identify the sender and receiver. A sender (i.e. a node on) creates a transaction by digitally signing it with a private key to transfer the value of from one address to another.

[2] INTEGRATING BLOCKCHAIN AND THE INTERPLANETARY FILE SYSTEM-

Author - Tudor Gabriel, Andrei Cornel.

This article explains the integration of blockchain and the Inter Planetary File System (IPFS), so it is important to understand the meanings of the terms. Blockchain is a distributed database that securely stores data in consecutive blocks, creating the encryption model of the chain by connecting these blocks together. Due to the encryption technology used in blockchain, it seems to eliminate the possibility of data being hacked. Digital data in businesses is published and distributed across many parts of the computer, making it transparent, decentralized, and resistant to network tampering. Since blockchain cannot store big data due to size limitation and high cost, an option is needed to store big data and enjoy the benefits of certificate. This can be used by distributed file storage systems such as the Inter Planetary File System (IPFS). It uses Inter Planetary File System (IPFS) point addresses. That is, check the contents of the archive and share it with all nodes to save data. Therefore, the Inter Planetary File System (IPFS) copies files over the network. Here is a simple table comparing data storage to Inter Planetary File System (IPFS). In the Interplanetary File System (IPFS), all computers are connected to each other according to a standard, and each user on the network has a copy of the downloaded file. Additionally, only users with the Interplanetary File System (IPFS) hash can access a copy of the file. Therefore, security is provided in the form of a hash, just like a password. Inter Planetary File System (IPFS) is called the "persistent web" because files always reside at a specific address; This, unlike classic Hypertext Transfer Protocol (HTTP), can prevent authentication access. for copies of information sent on peer to-peer (P2P) networks. Each copy has an encrypted hash that cannot be decrypted. Information is shared by these collaborators, called nodes, throughout the network. Each part of the network uses additional information to store relevant information. This will help you find the node that stores that content. Cryptographic hashes are needed to obtain all the information. Hashing splits and combines information from all different sources to create a downloadable file. Use Distributed Hash Table (DHT) to avoid data confusion here. Blockchain stores the hash address in a smart International Journal of Engineering Research and Technology (IJERT) contract. Ensure security by hiding content behind letters and numbers. Any change, regardless of motivation, will change the order of letters and numbers, indicating a discrepancy between the requested data and the data stored in the Interplanetary File System (IPFS).

[3] BLOCKCHAIN BASED DECENTRALIZED STORAGE SCHEME-

Author - Zuoting Ning and all.

The article aims to propose a management system based on blockchain technology that will be able to use the rest of the personal hard drives of users' worldwide storage sites, provide honest data for users after. The certificate is issued and verification ends, the user will pay the storage fee Lightning The supplier pays for its service through network technology. All certificates and payment information are stored on the blockchain to ensure the security and reliability of the system. Compared with the existing mainstream decentralized storage system, this solution is improved in the following aspects: System access to terms and payments. A decentralized storage solution based on blockchain. The user uploads the encrypted data to the broker, which sends the data to the storage provider and informs the user about the location where the data is stored. Once the integrity of the user and the storage provider is successfully verified, the user uses Lightning Network technology to pay the storage fee to the storage provider. Blockchain technology provides immutable information, only allowing new transactions but cannot modify or cancel transactions. However, data stored on the blockchain has a different cost model than stored data in terms

of size and value. For example, in terms of size, the Bitcoin blockchain provides an incredible amount of information on transactions. It is limited to 80 bytes and was reduced to 40 bytes in February 2014. Looking at the price, storing an 80-byte Bitcoin blockchain using Ethereum costs approximately $0.03617 and $0.007. As mentioned above, blockchain transactions only save small amounts of data, so it's important to choose what data to put on-chain and what data to buy off-chain. There are many file storage solutions designed to be blockchain friendly, such as Story, File Coin, Sia, and IPFS. This solution introduces the concept of peer-to-peer data distribution, where data is fragmented, encrypted and distributed across many nodes in the network to ensure that data is safe and available. The biggest problem in these systems is the lack of access control. These resources are important if blockchain-based applications equipped with data storage solutions are used in business and valuable areas such as finance and public works.

[4] SHARING SYSTEM USING BLOCKCHAIN NETWORKS-

Author - Mr. Nachiket More and all.

The purpose of this white paper is to complement our previous review of insecure data storage on Android smartphones by detailing threats and solutions. I think a full review of the Android storage model is necessary. So, we look at attacks, threats, and solutions to Android from 2013 to 2018. We also show a breakdown of Android storage threat models based on physical and software threats, with a variety of options for each group We address the problems. We are still looking for solutions to reduce all groups. That's why Android apps are designed to be securely stored and protected using blockchain and IPFS. Information technology has changed a lot in the last few years. Easy to use and easy to develop, is open source and has attracted the attention of developers who want to create content for the group. Cloud computing is considered the next big step in the use of all kinds of technology. From businesses to non-profit organizations to individual users, there are many programs that use cloud computing to deliver better, faster, and smarter computing. It seems so. This document aims to combine these two functions to create an Android cloud system and provide users with the experience of using the cloud. The performance of mobile devices, especially smartphones, has increased rapidly in recent years. Many users have high-performance smartphones and enjoy content on their smartphones for longer than on other devices. As a result, users are constantly changing content and the need to constantly manipulate data through searches has increased its importance. To overcome these problems, we have released a form sharing application for Android devices. We hope that the app will be a reliable and efficient data sharing tool for mobile devices.

[5] DESIGN OF DATA SHARING PLATFORM BASED ON BLOCKCHAIN AND IPFS TECHNOLOGY-

Author- Sanket Deshmukh and all.

This article focuses on the rapid development of information technology and digitalization, as well as the increasing amount of information that people create in their daily lives, including office information, photos and videos. With the development of blockchain technology, it now offers the following features: It brings a new approach to data security problems; i.e. "traceability", tampering" and "decentralization of power." Some scientists plan to build one. Data sharing solutions based on blockchain technology to ensure data privacy, security and sharing by accessing authorized locations and storing data. Data loss and to solve data tampering problems, some researchers have developed a data exchange platform combining blockchain and machine learning. Some researchers have proposed cloud + blockchain data chainless storage technology to achieve data transfer and storage.

PROPOSED SYSTEM

The proposed system for decentralized file storage and sharing prioritizes user control and transparency. Users can access content only with owner permission, while all actions are recorded on the Blockchain, ensuring transparency. Features such as logging, user access control, reputation systems, and hash verification facilitate monitoring and tracking of file-sharing activities.

Scalability, performance, and usability are improved through distributed storage (IPFS), optimized smart contracts, user-friendly interfaces, and performance benchmarking. Security and privacy are upheld with encryption, key management, anonymity, and continuous monitoring.

Furthermore, users can locate their data and track its history using hash values, enhancing transparency and accountability. Overall, the system aims to provide a secure, transparent, and efficient platform for decentralized file sharing while ensuring user privacy and data integrity.

IMPLEMENTATION

A. Adding a file

In the process facilitated by IPFS, each file uploaded by a node is broken down into chunks and structured as a Merkle Directed Acyclic Graph (DAG). Content IDs linked with the file are then recorded and passed to the permissions package. This package, in turn, registers the files in a smart contract using the add File function, consolidating essential transactions. Verification of transactions is achieved through the correct content IDs. If the transactions are successful, the IPFS code proceeds to store the blocks locally and registers as a content identifier provider. However, failure in a transaction prevents the ownership claim of at least one content identifier, consequently halting the storage and registration process of the entire file. In case an adversary node witnesses the transaction, they may attempt to claim the same content identification, acquiring the ability to set the content identifier. A smart contract may grant permission to it, although the success of the hostile node isn't guaranteed, as the first completed transaction prevails in case of competition. When transaction size limitations arise, multiple transactions may be utilized to register content identifiers. ACL-IPFS then monitors and reports on the overall success of this process.

B. Granting permission

Using the command line interface, ACL-IPFS enables users to easily grant access permissions to designated individuals. Subsequently, a transaction is generated and sent to a Blockchain node for processing. Henceforth, only the owner, identifiable by their metamask account, possesses the authority to grant or revoke permissions, as the content identity has been duly registered. This condition is enforced by a clause in the smart contract, as mentioned earlier.

C. Retrieving a file

If a node has been granted permission, it can request and obtain the chunks associated with files. It achieves this by finding suppliers of these chunks in the network and establishing connections with them to receive the necessary chunks for reconstructing the file. However, if authorization has not been granted or has been revoked, the request will not be fulfilled. Any individual connected to a metamask node has the ability to verify a user's rights for any content identifier and draw conclusions accordingly.

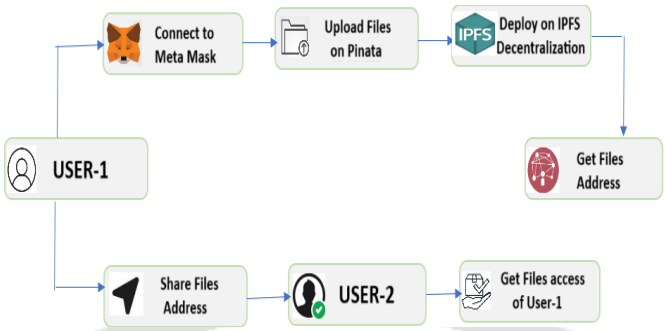
D. Connect to metamask wallet

To add a block to the Blockchain, users need to pay a gas fee, which is a small amount of cryptocurrency required for processing transactions. This fee is paid using a Blockchain wallet such as Ethereum, users need to install a browser extension and create an account. When a user uploads a file, a transaction is initiated, and the gas fee is automatically deducted from their Ethereum wallet. Since the system is operating on a development network (DevNet), users have the flexibility to add any amount of virtual currency to their wallet to ensure the smooth functioning of the project. This setup allows users to participate in the network and contribute to its operation without the need for real-world financial transactions.

E. Account creation

The number of files that can be uploaded to one account depends on the memory allocated by the developer. The user creates a new account within the application itself, when the limit is reached. Creating a new account involves generating a new key pair and assigning it to the account, which restricts access to previously uploaded files. The users can increase the memory allocation for a single account by contacting the developer. In this way users can continue uploading files without creating multiple accounts.

F. Architecture



CONCLUSION-

The integration of Blockchain into file-sharing and storage systems has revolutionized data management, primarily through its emphasis on decentralization. By distributing data across a network of nodes, Blockchain mitigates the vulnerabilities of centralized systems, such as single points of failure and susceptibility to tampering. This decentralized approach enhances security and transparency for users.

Ongoing research and development in Blockchain-based file sharing present numerous opportunities for innovation. Continued exploration of scalable, efficient, and user-friendly solutions, alongside integration with emerging technologies like artificial intelligence, holds promise for further advancements in decentralized file storage and sharing systems. Individuals to stay aware of these developments to contribute to the evolution of the field and harness the full potential of Blockchain for secure and accessible data management.

REFERENCES-

[1] Shalom, G. R., & Nirogi, G. R. (2022, September 30). Decentralized Cloud Storage Using Blockchain. International Journal for Research in Applied Science and Engineering Technology, 10(9), 1294–1300.
[2] Khalid, M. I., Ehsan, I., Al-Ani, A. K., Iqbal, J., Hussain, S., Ullah, S. S., & Nayab. (2023). A Comprehensive Survey on Blockchain Based Decentralized Storage Networks. IEEE Access11,10995-11015. https://doi.org/10.1109/access.2023.3240237
[3] Kumar B.R, M., & Ms. (2021, April). The Blockchain-Based Decentralized Approaches for Cloud Computing to Offer Enhanced Quality of Service in terms of Privacy Preservation and Security: A Review. IJCSNS International Journal of Computer Science and Network Security, 115(April 2021).
[4] Tao, J., & Ling, L. (2021). Practical Medical Files Sharing Scheme Based on Blockchain and Decentralized Attribute Based Encryption. IEEE Access,9,118771–118781. https://doi.org/10.1109/access.2021.3107591
[5] Wilkinson, Shawn, Jim Lowry, and Tome Bolsheviks. "Meta disk a blockchain-based decentralized file storage application." Story Labs Inc., Technical Report, hall (2014): 1-11.
[6] Ganache Private Ethereum blockchain environment: https://trufflesuite.com/docs/ganache/
[7] Decentralized File Storage (Interplanetary File System) using Blockchain by - Mr. Shobhit Khandare.
[8] Decentralized File Sharing using Blockchain Empowering Peer-to-Peer Collaboration: The Rise of Decentralized File Sharing By- Mr. Gaurav Jadhav.
[9] Cloud Storage. A comparison between centralized solutions versus decentralized cloud storage solutions using Blockchain technology By- Tudor Gabriel, Andrei Cornel.