Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 03, March-2023

# Decentralized File Storage (Interplanetary File System) using Blockchain

Jitesh Shamdasani
Department of Electronics and Telecommunication Engineering,
Vivekananda Education Society's Institute of Technology,
Mumbai,Maharashtra,India

Niraj Wadile
Department of Electronics and Telecommunication Engineering,
Vivekananda Education Society's Institute of Technology,
Mumbai,Maharashtra,India

Sanket Deshmukh
Department of Electronics and Telecommunication Engineering,
Vivekananda Education Society's Institute of Technology,
Mumbai,Maharashtra,India

Mobaashir Sayyed
Department of Electronics and Telecommunication Engineering,
Vivekananda Education Society's Institute of Technology,
Mumbai,Maharashtra,India

Mr.Shobhit Khandare
Assistant Professor,
Department of Electronics and Telecommunication Engineering,
Vivekananda Education Society's Institute of Technology,
Mumbai,Maharashtra,India

***Abstract:*** **Cloud storage is one of the leading options to store massive data, however, the centralized storage approach of cloud computing is not secure. On the other hand, Blockchain is a decentralized cloud storage system that ensures data security. Any computing node connected to the internet can join and form peers network thereby maximizing resource utilization. Blockchain is a distributed peer-to-peer system where each node in the network stores a copy of the blockchain thus making it immutable. This paper focuses on decentralized secure data storage, high availability of data, and efficient utilization of storage resources. When you add a file to IPFS, your file is split into smaller chunks, cryptographically hashed, and given a unique fingerprint called a content identifier (CID). This CID acts as a permanent record of your file as it exists at that point in time. When other nodes look up your file, they ask their peer nodes who are storing the content referenced by the file's CID. When they view or download your file, they cache a copy and become another provider of your content until their cache is cleared. IPFS uses content addressing to identify content by what's in it rather than by where it's located. Looking for an item by content is something you already do all the time. IPFS and many other distributed systems take advantage of a data structure called directed acyclic graphs**.

***Keywords: Blockchain, peer-to-peer system, IPFS, CID, cryptography, chunks, decentralized cloud storage, directed acyclic graphs.***

## I INTRODUCTION

Cloud storage is one of the main options for storing large amounts of data, but the cloud computing centralized storage approach is not secure. On the other hand, blockchain is a decentralized cloud storage system that ensures data security. Any computing node connected to the internet can join and form a peer-to-peer network, maximizing resource usage. A blockchain is a decentralized, peer-to-peer system where every node in the network keeps a copy of the blockchain, making it immutable. In the proposed system, users' files are encrypted and stored on multiple peers in the network using the IPFS (Interplanetary File System) protocol.

IPFS generates hash values. The hash value specifies the path to the file and is stored on the blockchain. This white paper focuses on distributed secure data storage, data high availability and efficient use of storage resources.

Large files cannot be stored efficiently on the blockchain. On the one hand, the blockchain is full of data that needs to be distributed on the blockchain network.

On the other hand, since blockchains are replicated across many nodes, they require a lot of storage space for no immediate purpose, especially if node operators don't need to see all the files stored on the blockchain. It also increases the cost of blockchain operating nodes as more data needs to be processed, transmitted and stored. IPFS is a file sharing system that can be used to store and share large files. It is based on a cryptographic hash that can be easily stored on the blockchain. However, IPFS does not allow users to share files with parties of their choosing.

Required if sensitive or personal data needs to be shared. Therefore, this article presents a modified version of the Interplanetary File System (IPFS) that uses Ethereum smart

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 03, March-2023

contracts to enable access-controlled file sharing. Smart contracts are used to maintain access control lists, and modified IPFS software enforces them. To do this, it interacts with the smart contract whenever a file is uploaded, uploaded or transmitted.
blockchain applications interact directly with a blockchain or smart contract to reach consensus on the execution of transactions, data or code.
A network of heterogeneous nodes stores blockchains, processes transactions, and executes smart contracts when needed. This causes the following problems when working with large data files: Normally, blockchain nodes don't need files to function, which leads to blockchain bloat, resulting in data being replicated across a large number of nodes.
On the one hand, storing large files on the blockchain is inefficient. Due to block size limitations, files must be split and reassembled off-chain.
Additional data relevant to reassembling files would also have to be stored, requiring either even more space or a distinct system that provides the reassembly information. If smart contracts are leveraged to directly store file parts, the data can more easily be accessed and the reassembly information could be stored as well. However, sending and storing large files, even partially, using smart contracts is expensive (for example regarding gas costs) and needs to be executed at every mining or verifying node portfolio.
On the other hand, operating the mining nodes becomes more expensive. More data needs to be propagated through the network, processed and stored by the node would thus require connections with higher bandwidths and more storage space to store the blockchain, even partially, thus leading to increased costs.

## III REVIEW OF LITERATURE

### *A. DECENTRALIZED CLOUD STORAGE USING BLOCKCHAIN*

This paper introduces a methodology that uses Blockchain applications to interact either directly with blockchains or with smart contracts in order to achieve consensus on transactions, data or execution. A network of heterogeneous nodes stores the blockchain, processes transactions and, if necessary, executes smart contracts. This leads to the following issue when working with large data files. Because the files are usually not required for the blockchain nodes to function, the blockchain becomes bloated, resulting in data being replicated on a large number of nodes data is analyzed.
The blockchain contains a series of transactions organized logically inside a block,these blocks are tied together using encryption in order to form a chain.
The technology itself forms a decentralized consensus allowing peers to reach consensus on the state of a transaction.
Information is usually stored in private databases maintained by each organization. The database technology is commonly described as the CAP theorem and relates to Consistency (C),Availability (A), and Partition 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.

### *B. INTEGRATING BLOCKCHAIN AND THE INTERPLANETARY FILE SYSTEM*

This paper introduces integrating blockchain and Interplanetary File System (IPFS), hence understanding what the terminologies mean is mandatory. Blockchain is a decentralized database that keeps a permanent record of information and stores it in blocks in consecutive, cryptographic patterns linking these blocks to form a chain. It almost nullifies the chances of hacking the database due to the cryptographic encryption technique used in the blockchain. The digital ledger
transaction is duplicated and spread across various nodes of the computer which makes it transparent, decentralized, and immune to tampering network. Since blockchain cannot store large files due to size limitations and high cost, an alternative is necessary to store large files and get the benefits of the decentralized database.
This can be achieved with distributed file storage systems such as the Interplanetary File System (IPFS). Interplanetary File System (IPFS) uses content addressing. That is, check what is in the file and share it with all nodes to save the file. Therefore, the Interplanetary File System (IPFS) copies files over the network.
Here is a simple chart comparing a traditional file storage system to the Interplanetary File System (IPFS). In the Interplanetary File System (IPFS), all computers are interconnected according to the standard, and all users on the network have copies of downloaded files. Additionally, copies of the files are accessible only to those with the Interplanetary File System (IPFS) hash. So, providing security as a hash, acts like a password. The Interplanetary File System (IPFS) is called the "Persistent Web" because files always reside at a specific address, unlike the classic
Hypertext Transfer Protocol (HTTP), which protects against unauthorized access.Copies of files are distributed over a peer-to-peer (P2P) network. Each copy has a cryptographic hash of that cannot be decrypted. Data is shared across the network from these peer nodes, called nodes. Each node in the network uses extra information indexing to store relevant content. This will help you find out which node stores which content needs these cryptographic hashes to get complete data. The hash receives data from all different nodes where the data is split and combined to form the downloaded file. Data scrambling here is prevented using a Distributed Hash Table (DHT). Blockchain stores hash addresses in smart

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 03, March-2023

contracts. Increase security by hiding the contents of file behind a series of letters and numbers.
Any change, regardless of motive, completely changes this sequence of letters and numbers , which proves the lack of consistency between the requested file and the file stored on the Interplanetary File System (IPFS).

*C. BLOCKCHAIN BASED DECENTRALIZED STORAGE SCHEME*

The paper intends to propose a decentralized storage system based on blockchain technology, which can make full use of the remaining space of personal hard disks of users around the world.The storage provider performs a data integrity certificate for the user, and after verifying that the verification is passed, the user pays the storage fee to the storage provider through lightning network technology. All proofs and payment information are stored in the blockchain, which
guarantees the security and credibility of the system. Compared with the current mainstream distributed storage systems this scheme has been improved in terms of system access and payment method. a distributed storage scheme based on blockchain. The user uploads the encrypted data to the middleman, and the middleman sends the data to the storage provider and informs the user of the data storage location. After the data integrity certificate is completed between the user and the storage provider, the user uses lightning network technology to pay the storage fee to the storage provider.
Blockchain technology provides immutable data storage in that it allows transactions to only be appended and never to be modified or removed. However, the data storage on the blockchain has a cost model that differs from conventional data storage in terms of size and cost. For instance, in terms of size, the Bitcoin blockchain provides the store of arbitrary data in transactions. It was limited to80 bytes which has been reduced to 40 in February 2014. In terms of cost, storing 80 bytes on the Bitcoin blockchain costs roughly US$0.03617 and US$0.007 when using Ethereum.As described, blockchain transactions hold only very small amounts of data, wherefore it is crucial to choose what data should be placed On-chain and what should be kept Off-chain. There are several Off-chain data storage solutions designed to be friendly to blockchains, such as Storj,FileCoin, Sia and IPFS. These solutions share the idea of a peer-to-peer distributed file system, where the data is shredded, encrypted and distributed to multiple nodes in the network to ensure their safety and availability. One main issue with these systems is the lack of access control.These features are extremely important if blockchain-based applications equipped with off-chain data storage solutions are deployed in operational and sensitive environments such as the financial and public sectors.

*D. SHARING SYSTEM USING BLOCKCHAIN NETWORKS*

This white paper is intended to supplement our previous review of insecure data storage on Android smartphones by expanding the range of security threats and solutions. I think thorough testing of the Android storage model is warranted. So, we are looking at attacks, threats and solutions for Android from 2013-2018. We also propose a phased separation of the threat model for Android data storage based on physical and software threats, and consider several issues for each class .We are also exploring solutions to reduce each category of.Therefore, the Android app is designed to securely store and secure using BlockChain and IPFS.There have been dramatic changes in information technology over the past few years. Easy to use and easy to develop, the is open source and has caught the attention of developers who want to create content for the masses. Cloud computing is known as the next big step in all its forms using standard technology. From businesses to non-profit organizations and individual users,has many programs that can use cloud computing to provide better, faster and smarter computing. Seems to.This document aims to combine these two tasks to create an Android cloud system and provide users with a cloud computing experience in the palm of their hands. The performance of
mobile devices, most of which are smartphones, has improved rapidly in recent years. Many users have high-performance smartphones and enjoy content longer on their smartphones than on other devices. As a result users are constantly exchanging content, and the requirements for exchanging files through the extension call have increased significantly. To overcome these problems, we introduce seamless file sharing app for Android devices.We expect the proposed application to be a reliable and cost-effective file sharing solution for mobile devices.

*E. DESIGN OF DATA SHARING PLATFORM BASED ON BLOCKCHAIN AND IPFS TECHNOLOGY*

This essay focuses on the rapid development of information technology and digitalization, as well as the rising amount of data produced by people in their daily lives, including office documents, photos, and videos. With the advancement of blockchain technology, it now exhibits the following features:Providing a fresh approach to the issue of data security is "traceability," "hard to tamper with," and "decentralisation." A data sharing plan based on blockchain technology has been presented by certain academics to guarantee data privacy security and sharing through access permission setting and data storage. To address issues like data loss and data tampering, some academics have created a platform for exchanging data that combines blockchain and machine learning technology. The data unchained storage mechanism of cloud + blockchain has been suggested by certain academics to realise data transfer and storage.

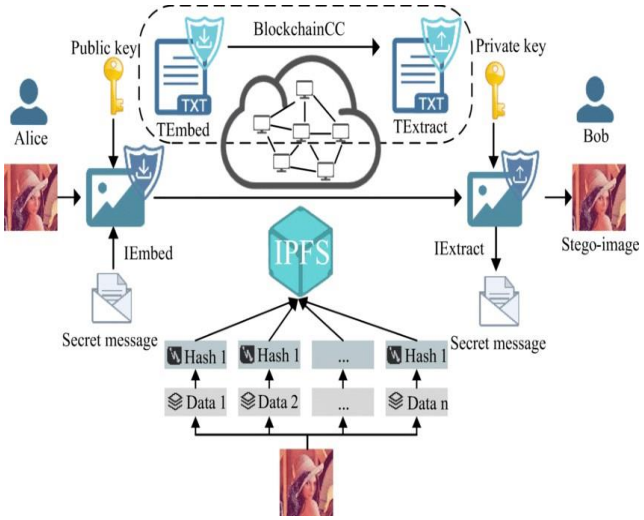
## IV PROPOSED PROJECT SYSTEM

With the ongoing development of information technology, people's need for computing and resource storage has recently demonstrated a tendency of rapid rise. In order to satisfy their desire for greater computational power and larger storage, people are continually investigating novel methods of storage room Cloud computing has drawn the attention of academic and commercial circles since the introduction of computing paradigms like peer-to-peer, grid,

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 03, March-2023

utility, and distributed computing. As a result of the way cloud computing divides up processing work among huge groups of computers, different application systems can access computational resources, storage capacity, and software services as needed. Blockchain technology was presented by Nakamoto in 2008 when he released "Bitcoin: a peer-to-peer electronic cash system," and it was built and used in Bitcoin.The use of blockchain technology.

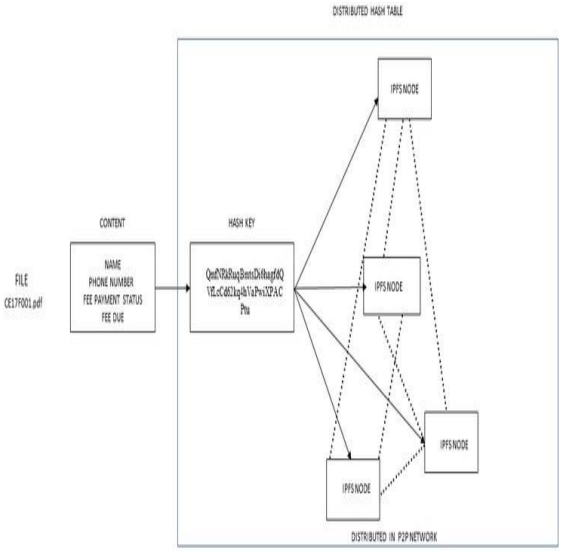
### *A. BLOCK DIAGRAM*




Blockchain-IPFS Architecture

Whenever a node uploads a file, as per IPFS the node creates chunks and the associated Merkle DAG.During this process all content identifiers corresponding to thefile are collected. These are then passed on to the permissions package, which registers the files in the smart contract, using the addBlock functions. This is done by forming the trans-actions necessary for this purpose the permissions systemfurthermore allows the verification of transactions based on the corresponding content identifiers. If the transactions succeed,the execution returns to the IPFS code which stores the blockchain the local storage and then registers itself as a provider of the content identifiers. If the transaction fails, ownership of atleast one of the content identifiers and therefore the whole file could not be claimed.

### B. FILE STORAGE IN IPFS



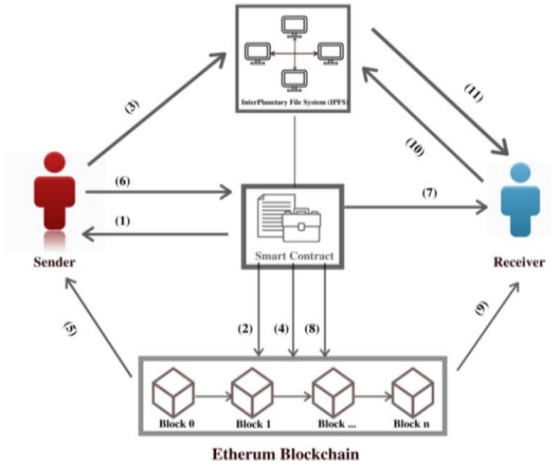

File Storage in IPFS Block Diagram

The copy of the file is distributed over the peer-to-peer (P2P) network. Every copy has a cryptographic hash that cannot be decoded. Data will be shared across a network on these peers called nodes. Each node in the network stores the content whose relevancy it finds with the help of additional indexing of information. It helps to find which node is storing what content. To retrieve the full data, there is a need for these cryptographic hashes. The hash gets data from all the other nodes throughout which data was split and combine to form the file, that was uploaded. Here data scrambling is avoided by a distributed hash table (DHT). The blockchain stores hash address in the smart contract. It strengthens the security by hiding the content of the file within a series of letters and digits. Any alteration regardless of the motive changes this series of letters and digits completely which gives proof that there is no consistency between the file requested and the file that is stored in Interplanetary File System (IPFS).

### C. WORKING

The single point of failure problem of relying on third party services is inexorably carried over by the conventional mechanism of centralised data processing systems. In many instances, data security is jeopardised because cloud storage service providers must endure pointless conflicts like political censorship. Additionally, individuals might not be able to access their own data as a result. These data indicate that, Decentralized storage methods will be necessary in the future to offer people services for sharing and storing data. This approach suggests decentralised data redundancy, which will guarantee that copies of the data are kept on each node in a peer-to-peer network. A distributed peer-to-peer system known as blockchain uses each network node to exchange data.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 03, March-2023




Working Diagram of Blockchain-IPFS

Hashes are generated by IPFS. The blockchain stores the hash value, which identifies the file's path. This project is focused on decentralised secure data storage, high data availability, and effective storage resource usage. Your file is divided into smaller pieces, cryptographically encrypted, and given a special fingerprint known as a content identifier when you add it to IPFS (CID). Your file as it was at that time is permanently preserved by this CID. The peer nodes that are storing the information referred to by the CID of the file are consulted by other nodes when they look up your file. They make a copy of your material when they view or download it, and until their cache is empty, they serve as another distributor of your content.

## D. RESULTS

The IPFS module, the quick retrieval module, the encryption and decryption module, and the blockchain module make up the data sharing platform system.

The browser serves as the system's user interface, allowing users to interact with each module through web pages and send HTTP requests to the background to carry out the appropriate tasks.

In order to create interactivity, data is converted into pages using technologies like HTML, CSS, and others. Blockchain, IPFS, and data are provided by the front end of the background.

Other features include quick data retrieval, encryption and decryption.

Realization of a data-sharing platform: Users upload data through a browser; data input is completed by a background blockchain module; for encrypted or decrypted files, a background encryption and decryption module completes the data input.

## E. REFERENCES

[1] Meet Shah, Grinal Tuscano,"Decentralized Cloud Storage Using Blockchain", 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)

[2] Reema Chandan Roychaudhary, Mercy Gill,"Integrating Blockchain and the InterplanetaryFile System", 2021 Journal of Interdisciplinary Cycle Research

[3] Yan Zhu, Chunli Lv, Zichuan Zeng "Blockchain-based Decentralized storage scheme",2019 Journal of Physics:Conference Series,Volume 1237,Issue 4

[4] Ranjeet Deshmukh, Rajashri Sadafule,Mihir Nevpurkar, "Sharing System using Blockchain and IPFS",2020 International Journal of Injury Control and Safety Promotion

[5] Weijing Li, Zicheng Zhou, Wen Fan,Juan Gao,"Design of Data Sharing Platform based on Blockchain and IPFS Technology",2022 Wireless Communication and Mobile Computing

[6] Ipfs.tech,'A peer-to-peer hypermedia protocol',2022[online],Available:https://docs.ipfs.tech/

[7] Wikipedia.,'Blockchain',2022.[Online] Available:https://en.wikipedia.org/wiki/Blockchain

[8] Ganache-Private Etherum blockchain environment:https://trufflesuite.com/docs/ganache/