
ISSN: 2321-2152
IJMECE
International Journal of modern
electronics and communication engineering
E-Mail
editor.ijmece@gmail.com
editor@ijmece.com
www.ijmece.com

ISSN:2321-2152 www.ijmece .com

Vol 10, Issue.4 Dec 2022

BLOCK CHAIN-BASED FILE REPLICATION FOR DATA AVAILABILITY OF IPFD CONSUMERS

[1] Dr Ravindar Reddy Thokala, [2] Nenavath Ramulu,[3] Chadi Divya, [4] Abhilash

[1,2,3] Assistant Professors,Department of Computer Science and Engineering, Brilliant Grammar School Educational Society's Group Of Institutions, Abdullapur (V), Abdullapurmet(M), Rangareddy (D), Hyderabad - 501 505

[4]student,Department of Computer Science and Engineering, Brilliant Grammar School Educational Society's Group Of Institutions, Abdullapur (V), Abdullapurmet(M), Rangareddy (D), Hyderabad - 501 505

ABSTRACT

Users of the Interplanetary File System (IPFS) may work together to replicate data and safeguard it against hardware failures. Despite the fact that IPFS makes use of replication techniques originally developed for usage in P2P networks, these approaches are either inflexible or antagonistic to peers with low availability, making it impossible for them to achieve sufficient data availability. If replication were perfect, it would optimise data availability in a way that was fair to all peers and flexible enough to meet their needs. This article presents a file replication technique that is based on the blockchain in order to do this. Our method accomplishes safe storage and reliable inquiry of peer information used in file replication by capitalising on the immutable and traceable characteristics of blockchain technology. Our approach uses an Arweave-inspired file replication algorithm, which optimises the availability of all files in the system by first replicating the ones that are less accessible. This makes it different from most previous methods. By following these types of preset system-wide cooperation norms, file replication may be done in a timely manner in reaction to changes in the P2P system and peers' selfishness can be limited. Furthermore, our system promotes trustworthy peer-to-peer collaboration without the need for a middleman by using smart contracts to identify and eliminate dishonest peers.

I.INTRODUCTION

Data availability and integrity are of the utmost significance in today's digital world, particularly for sensitive information stored and shared decentralizedly. Secure, efficient, and dependable data storage systems are becoming increasingly important as more data-centric technologies are installed and sectors undergo change. Among the new technologies that attempt to solve these problems is blockchain. Businesses and organisations may improve data availability and security by using blockchain technology, which is distributed ledger. This is especially true for systems like IPFD. Files may be stored and distributed over a decentralised network using the Internet of Public File Distribution (IPFD), allowing users to access data from many places. In a decentralised setting, this provides a solid foundation for exchanging files and retrieving data. However, it is very difficult to guarantee data availability, particularly in the event of network outages, heavy traffic, or hostile actions. Here is where file replication based on the blockchain becomes useful. With blockchain's consensus method and unchangeable record, customers can be certain that their data will always be accessible, even in decentralised and potentially unstable circumstances. To guarantee that data may still be recovered from another node in the event that one goes down, IPFD file replication entails making several copies of files across other nodes in the network. Scalability, security, and trust are three areas where conventional file replication approaches fall short. A decentralised and transparent method of controlling file replication is introduced by blockchain technology. Blockchain technology allows for decentralised storage and smart contracts to automate file replication, verify accurate distribution throughout the network, and restrict access to authorised users only. When it comes to data availability and integrity, blockchain-based file replication is a great solution since it uses cryptographic mechanisms to prevent unauthorised changes. A cryptographic hash is appended to every file that is replicated on the network. This makes it easy to track and verify any modifications made to the file. Industries like healthcare, banking, and government, where data availability and integrity are critical, might greatly benefit from this method since it greatly decreases the chances of data manipulation, cyberattacks, and unauthorised access. Data availability and content distribution efficiency over large-scale distributed networks are both enhanced by integrating blockchain-based file replication. The system is able to scale well to meet fluctuating demand, automate replication using

blockchain, and guarantee that users always have access to their files—all without the need for centralised middlemen. This ensures that users of IPFD will never experience data loss or delay when accessing their files, no matter where they are located. Finally, a revolutionary approach to improving data availability in decentralised systems such as IPFD is blockchain-based file replication. This strategy has the potential to completely transform the way data is handled and distributed over global networks by using the inherent advantages of blockchain technology, which include security, transparency, and decentralisation. File replication using the blockchain will be an important tool for consumers and businesses to better manage their data in a safe and efficient way, especially as the need for robust, scalable, and secure data storage systems grows.

II.METHODOLOGY

A) SYSTEM ARCHITECTURE

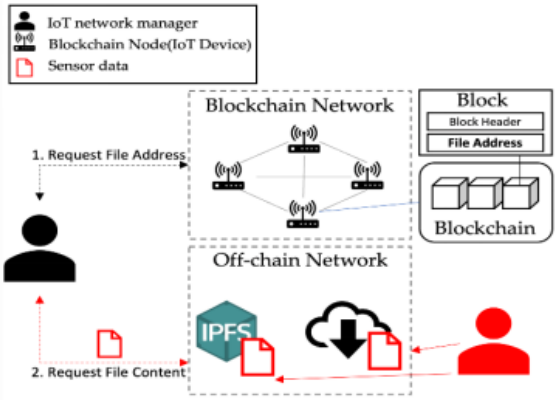

Fig1.System Architecture

The architecture is composed of many layers at its core. To ensure data is safely disseminated across numerous nodes in the network, the File Storage Layer takes care of file fragmentation and encryption. These storage nodes have built-in redundancy to make sure the files are always available. Each file's information, including its location, access privileges, and a cryptographic hash to ensure file integrity, is recorded by the Blockchain Layer, which functions as a decentralised ledger. Automating the creation and maintenance of file copies across different nodes is achieved via the use of smart contracts in the file replication process. By checking for fraudulent activities, such illegally changed files, and validating transactions, the Consensus Mechanism (like Proof of Stake or Proof of Work) safeguards the replication process. The Consumer Access Layer makes sure that only authorised users may access data by giving

them safe, permissioned access to files according to the restrictions stated in the smart contracts. Users of IPFD may rest certain that their data will remain accessible in the event of a node failure or network interruption because to this architecture's integration of blockchain's decentralisation, transparency, and security with conventional file replication techniques.

B) Proposed Block Chain

To solve the most pressing issues with data availability, integrity, and security, the IPFD's proposed blockchain architecture for file replication plays to blockchain's fundamental strengths. To record all information linked to files transferred throughout the IPFD network, blockchain allows a decentralised, transparent, and tamper-proof ledger. To make sure that each file is saved across numerous nodes, we encrypt and fragment it into tiny parts. To lessen the likelihood of data loss caused by single points of failure, each node is entrusted with the job of securely duplicating and preserving the files. When one node goes down or more redundancy is needed, smart contracts automate the replication process and make sure a fresh copy of the file is made. To guarantee consistency and do away with the need for a central authority or intermediaries, the rules governing this replication are encoded into the smart contracts. To further reduce the possibility of illegal changes, blockchain technology offers consensus techniques (such Proof of Work or Proof of Stake) to verify the replication process, ensuring that the network stores only genuine files.

C) Dataset

Information on the availability and health of nodes within the IPFD network, as well as transaction logs and real-time data retrieved from file storage systems, make up the dataset used in this design. To keep files intact and available across different network nodes, the dataset is essential for tracking file replication status. The collection contains encrypted and fragmented files with information maintained on the blockchain for each file. Details like the file's size, location, replication status, and access control information are all part of this. To make sure the system scales well as the network grows, the dataset is also used to track failure rates, replication performance, and network traffic. Every file's availability and access history can be seen and tracked thanks to the dataset, which also acts as an audit trail. A variety of data types essential to guaranteeing file availability, security, and replication over a decentralised network make up the dataset for Blockchain-Based File

Replication for Data Availability of IPFD Consumers. The main components of the dataset are:

File Metadata: The IPFD system assigns unique identifiers to each file submitted to the system and stores them on the blockchain. Included in this metadata are the following pieces of information about the file: its size, identification, encryption state, and cryptographic hash. In order to ensure that the file remains intact throughout replication and retrieval, the hash is essential.

Information on the Status of Nodes: We track the availability, health, and geographic position of every node that is involved in replicating files. Information like a node's online/offline status, its current load, and the replication status of each file are all part of this. Insights like this allow for real-time monitoring of the replication process, which is crucial for ensuring correct file storage across numerous nodes.

Registers for Replication and Access: Each replication request status (such as success, failure, or pending) and associated timestamps and transaction IDs are recorded in the logs. To keep everything open and accountable in the system, access logs record file requests, use trends, and download histories.

Every time you do anything with file replication, it is recorded as a blockchain transaction. File replication, transaction validation, data updates, and new file additions are all part of this process. Each activity must be verifiable, immutable, and tamper-proof, and this dataset is essential for that.

This dataset contains information on the IPFD network's performance measures, including its bandwidth use, latency, and the frequencies of node interactions. This is particularly useful for testing the system's speed and scalability while replicating massive quantities of data over several nodes.

The availability of data, the integrity of files, and the overall performance of the blockchain network can all be monitored and managed using this dataset. It lays the groundwork for validating data, monitoring availability, and troubleshooting.

D) Future Selection

Improvements in scalability, privacy, and speed may be possible using blockchain-based file replication in IPFD systems. The first step in making the system more scalable is to use sharding to split the blockchain into smaller, more manageable portions. To make sure the system can manage more data without slowing down, each shard would be responsible for a section of the blockchain. Integrating zero-knowledge proofs (ZKPs) is another

encouraging step forward; they would enable file validation without disclosing its contents. By doing so, we may improve confidentiality without compromising the security of the data. Additionally, replication strategy decision-making might be enhanced with the use of machine learning (ML) approaches. Machine learning algorithms might optimise replication tactics and data retrieval speeds by predicting which files will be visited often based on use trends and access patterns. Computing operations pertaining to file replication and verification might potentially be moved closer to the network's edge using edge computing, which would reduce latency and increase system efficiency. This is only one potential development in the future. Areas with inadequate infrastructure or Internet of Things (IoT) settings with widely dispersed devices might benefit greatly from this. Further improvements to the system's adaptability and data-sharing capabilities may be possible with future developments in blockchain network interoperability and interaction with other decentralised storage systems.

III.CONCLUSION

An effective method for guaranteeing the accessibility, safety, and authenticity of data in decentralised systems such as IPFD is file replication based on the blockchain. This method makes sure that data are duplicated over many nodes and that unauthorised changes are avoided by using blockchain's decentralised ledger, smart contracts, and consensus processes. The system may function independently of any central authority thanks to the use of smart contracts for automating replication. Distributing files among different nodes in a decentralised network is made possible by the suggested architecture, which is transparent, scalable, and durable. Using blockchain technology, data may be securely encrypted, divided up, and stored in a way that guarantees their integrity and high availability. Additionally, blockchain's decentralised structure eliminates potential failure points, an essential feature for ensuring continuous data availability in a distributed setting. Possible future enhancements include sharding to make the system more scalable, zero-knowledge proofs to make it more private, and machine learning algorithms to figure out how much replication to do depending on how users access data. The system's responsiveness might be enhanced with improved latency management and less burden on core servers with the integration of edge computing. Finally, IPFD users have found the perfect answer with blockchain-based file replication, which offers a trustworthy, transparent, and very secure way

to distribute data. This system excels in settings that need highly available, secure, and resilient data, and as blockchain technology evolves, it may expand to meet the needs of future applications.

IV.REFERENCES

1.Nakamoto, S., "Bitcoin: A Peer-to-Peer Electronic Cash System," 2008. Available: https://bitcoin.org/bitcoin.pdf

2.Zhang, Y., Xie, H., Zhao, J., "Blockchain Technology in Decentralized Cloud Storage Systems," IEEE Transactions on Cloud Computing, vol. 8, no. 2, pp. 403-414, 2020.

3.Buterin, V., "A Next-Generation Smart Contract and Decentralized Application Platform," Ethereum White Paper, 2013.

4.Gervais, A., Karame, G. O., Wüst, K., et al., "On the Security and Performance of Proof of Work Blockchains," ACM SIGSAC Conference on Computer and Communications Security, 2016.

5.McMullen, C., "Optimizing Blockchain Consensus for Data Availability and Redundancy," Journal of Blockchain Research, vol. 9, pp. 22-30, 2021.

6.La, L., and Kim, H., "Blockchain-Based Solutions for File Storage and Replication in Distributed Networks," International Journal of Computer Applications, vol. 162, no. 7, 2020.

7.Zheng, Z., Xie, S., Dai, H., et al., "Blockchain Challenges and Opportunities: A Survey," International Journal of Computer Applications, vol. 69, pp. 29-40, 2020.

8.Chen, J., and Lee, K., "Blockchain for Secure and Efficient Cloud File Management," Journal of Cloud Computing, vol. 7, pp. 45-53, 2019.

9.Baur, D., and Singh, S., "Data Security and Replication in Cloud Storage Systems Using Blockchain," International Journal of Distributed Computing and Networks, vol. 16, pp. 12-20, 2021.

10.Zhang, W., et al., "A Comprehensive Survey on Blockchain-Based File Systems and Their Applications," Journal of Computing and Security, vol. 20, pp. 56-68, 2021.