Technoarete Transactions on Advances in Computer Applications (TTACA)
Vol-2, Issue-3, September 2023
e-ISSN: 2583-3472

Anonymous Peer-To-Peer Distributed File System

Dr. S.T.Patil[1], Prakhar Rai[2]*, Swaraj Puri[3], Preeti Wagh[4]

[1] Professor, Vishwakarma Institute of Technology, Pune, India
[2][3][4] Vishwakarma Institute of Technology, Pune, India
*Corresponding Author Email: prakhar.rai20@vit.edu

Abstract

The rise of decentralized technologies has brought about significant advancements in the field of data storage and sharing especially a need for handling data in the format of files. This research paper explores the design and implementation of an anonymous peer-to-peer distributed file system (AP2P-DFS) that leverages blockchain and the InterPlanetary File System (IPFS). This system aims to provide secure, censorship-resistant, and privacy-centric file storage and sharing capabilities while maintaining user anonymity. The paper discusses the architectural design, key components, and security features like access control and anonymity of the AP2P-DFS and presents experimental results to evaluate its performance and effectiveness.

Keywords

Access Control, Blockchain, Data protection, Decentralization, InterPlanetary File System (IPFS)

INTRODUCTION

Blockchain technology is a revolutionary and secure way of storing and transmitting digital information. At the heart of blockchain's functionality are cryptographic keys, which are instrumental in ensuring the integrity and security of data. A blockchain is a decentralized and distributed ledger that records transactions across a network of computers, known as nodes. Each participant in the network possesses a pair of cryptographic keys: a public key and a private key. [1]

The public key serves as an address, much like an email address, through which others can send digital assets or verify your identity within the blockchain network. However, it cannot be used to reverse-engineer the private key, ensuring the security of your assets. On the other hand, the private key is a closely guarded secret and is used to sign transactions, proving your ownership of digital assets and enabling you to transfer them securely. When a transaction is initiated, the private key is used to create a digital signature, which is a unique identifier for the transaction. This signature is then verified by the network using the corresponding public key, ensuring the authenticity of the transaction. [2]

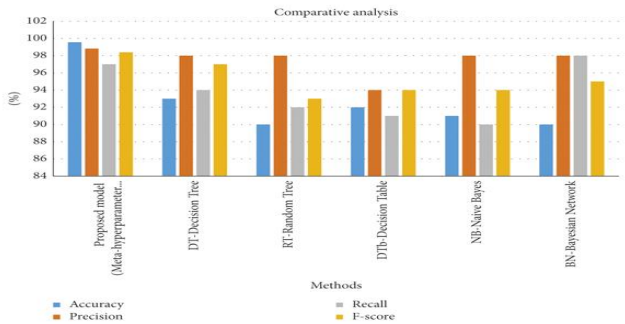

Figure. 1. Comparative analysis of IPFS and Blockchain based on factors like Accuracy, Precision, Recall and F-score

Blockchain technology operates on a principle of consensus among its nodes, and this is where the cryptographic keys come into play. In order for a transaction to be added to the blockchain, it must be verified by a majority of the nodes on the network. This verification process ensures that the transaction is legitimate and not the result of fraud [3]. Once consensus is reached, the transaction is recorded as a new block on the blockchain, creating an immutable and transparent history of transactions. [4]

Furthermore, blockchain technology offers a secure and efficient way to handle digital data through its integration with the InterPlanetary File System (IPFS). IPFS is a decentralized and peer-to-peer protocol that allows the storage and retrieval of data in a distributed manner. Unlike traditional centralized data storage systems, IPFS does not rely on a single server or data center to store information. Instead, it breaks data into smaller chunks and distributes them across a network of nodes. [5]

When blockchain is combined with IPFS, it offers an innovative solution to the problem of data permanence and availability. In this setup, the blockchain stores references, or content addresses, to the data stored on the IPFS network. These references are hashed representations of the data, making it highly tamper-resistant [6]. When a user or node wants to retrieve a specific piece of data, they use the reference stored on the blockchain to locate the content on the IPFS network. [7]

This integration of blockchain and IPFS has several benefits. Firstly, it enhances data security and integrity, as data on IPFS is highly redundant and resistant to censorship. Additionally, it promotes decentralization, as data is not stored on a single, vulnerable server but rather distributed across the network. This ensures that data remains accessible even if individual nodes go offline. Moreover, it increases the efficiency of data retrieval since users can access data from the nearest and fastest source on the network, reducing

e-ISSN: 2583-3472

latency and improving overall performance. [8]

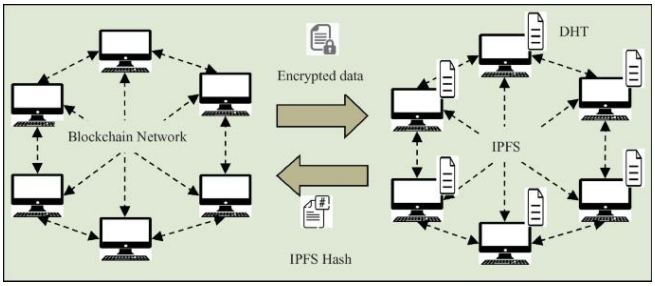

Figure. 2. Generic system design of IPFS and Blockchain network using encrypted data and Hash

In conclusion, blockchain technology relies on cryptographic keys to ensure the security and authenticity of transactions. Public and private keys work together to create a secure and transparent ledger, which is maintained by a decentralized network of nodes [9]. The integration of blockchain with IPFS further enhances data security, decentralization, and efficiency by storing data references on the blockchain and the actual data on a distributed and resilient IPFS network. This combination of technologies has the potential to transform the way we store and share digital information, making it more secure, transparent, and accessible to all. [10]

The proliferation of centralized data storage solutions has raised concerns regarding data privacy, security, and censorship. Traditional centralized systems are vulnerable to data breaches, government surveillance, and single points of failure. Decentralized technologies like blockchain and IPFS offer a promising solution to these issues by enabling the creation of a distributed file system that is secure, anonymous, and resistant to censorship. [11]

In this paper, we present an Anonymous Peer-to-Peer Distributed File System (AP2P-DFS) that combines the advantages of blockchain technology and the InterPlanetary File System (IPFS) to achieve a secure and anonymous data storage and sharing platform. The AP2P-DFS ensures that users can store and retrieve files without revealing their identity, and it prevents unauthorized access and censorship.

Background

Blockchain Technology Blockchain is a distributed ledger technology that ensures immutability and transparency of transactions. It employs cryptographic techniques to secure data and achieve consensus among participants. In our AP2P-DFS, we utilize a blockchain to manage user identities, access control, and record file metadata.

InterPlanetary File System (IPFS) IPFS is a decentralized and distributed file system that employs a content-addressable approach to file storage. Files are referenced by their cryptographic hashes, making data retrieval efficient and resistant to censorship. We incorporate IPFS into the AP2P-DFS to handle the actual storage and retrieval of files.

LITERATURE SURVEY

A. Secure and Transparent Election System for India using Block chain Technology

Authors find that integrating blockchain technology into the current voting system can solve issues with data accessibility, scalability, dependability, and most crucially security.

These characteristics are highlighted by explaining how blockchain is a widely distributed system that results in high data availability. When a block is added to the blockchain, it signifies that the largest possible number of network peers have authenticated and confirmed it. This feature makes the system decentralized and more dependable.

B. An Innovative IPFS-Based Storage Model For Blockchain (2018)

A continual growth in data volume caused by blockchain's capacity to only upload data rather than erase it restricts the number of nodes that can join the network. Including IPFS into a blockchain storage model system is recommended by this study. where, after depositing the transactions into the IPFS network, the miner adds the returned IPFS hash of the transaction into a block. It has drastically cut down on blockchain storage

C. Content Addressed P2P File System for the Web with Blockchain-Based Meta-Data Integrity (2019)

The limitations of the HTTP protocol are beginning to be seen as more content is being copied online. Files can be retrieved from a single source using this protocol. A peer-to-peer file system called IPFS makes it possible to store files in a decentralized setting. IPFS can handle massive volumes of data thanks to the immutable and permanent links and meta-data that are stored as Blockchain transactions. This timestamps and secures the data rather than requiring it to be placed on the chain itself. This article presents an architecture that combines the Blockchain's integrity preservation capabilities with IPFS's decentralized file storage system in order to store and distribute data on the Web.

Table. 1. Analysis of research papers based on IPFS, Blockchain and Decentralization control

Paper No.	Blockchain	IPFS	IPFS + Blockchain	Decentralized control over file access
[1]	✕	✕	✕	✕
[2]	✓	✕	✕	✕
[3]	✓	✕	✕	✕
[4]	✓	✕	✓	✕
[5]	✓	✕	✓	✕
[6]	✓	✕	✓	✕
[7]	✓	✓	✕	✕
[8]	✓	✕	✕	✕
[9]	✓	✕	✓	✕
[10]	✓	✓	✕	✕
[11]	✓	✕	✕	✕
Our work	✓	✓	✓	✓

Technoarete Transactions on Advances in Computer Applications (TTACA)
Vol-2, Issue-3, September 2023
e-ISSN: 2583-3472

METHODOLOGY

System Components:

1. **Tor**
2. **User Interface**
3. **IPFS**
4. **Blockchain**

Tor: The Tor network is designed to provide anonymity and privacy for internet users by routing their internet traffic through a series of volunteer-operated servers (nodes) to obfuscate the origin of the traffic. Tor clients enable users to browse the internet, send messages, and access online resources while masking their IP addresses and enhancing online privacy. Therefore, the user interface will have a tor wrapper which will provide user anonymity and security.

User Interface(UI): It reads a QR that contains user identity (wallet key), data encryption, and decryption keys. Using UI, the user selects a file and enters the list of receiver addresses to which access needs to be given. The system will proceed to encrypt a file with an AES key, which will be further encrypted using the public key based on the address of each receiver.

Encrypted keys along with respective addresses are entered in an encrypted file. File is sent to the IPFS node and the content ID is received(CID), this is sent to the blockchain network.

Using a UI user can access the files, and based on the wallet key user can gain accessible files CID. Based on CID, user downloads a file from IPFS, system searches the file for the user's address based on which the encrypted key is fetched from the file. The encrypted key is decrypted using the user's private key which returns the AES key to decrypt the file. Based on the AES key file will be decrypted. If an unauthorized user downloads the file, the system will not be able to decrypt it as the AES Key is not encrypted using the unauthorized user's public key, the decryption process using its private key will not work.

IPFS: IPFS, or the InterPlanetary File System, is a decentralized protocol for storing and sharing information on the internet. It uses a content-addressable system, making data retrieval more resilient and efficient by referencing content based on cryptographic hashes rather than traditional URLs. IPFS aims to create a more decentralized and robust web infrastructure.

Blockchain: It is a decentralized and immutable digital ledger technology used for secure and transparent record-keeping of transactions across a network of computers. This component maintenance two chains-

Chain1: stores ledger which maps the user's wallet address to the user's public key(used for data encryption).

Chain2: stores ledger containing - content ID (CID), address of the user who writes a file, and addresses of authorized receiver users.

The system on the UI side requests chain1 to gain access to authorized user's public key, and when the file is stored in IPFS, a storage request is made on chain2 to record file's content ID and list of authorized users.

IMPLEMENTATION

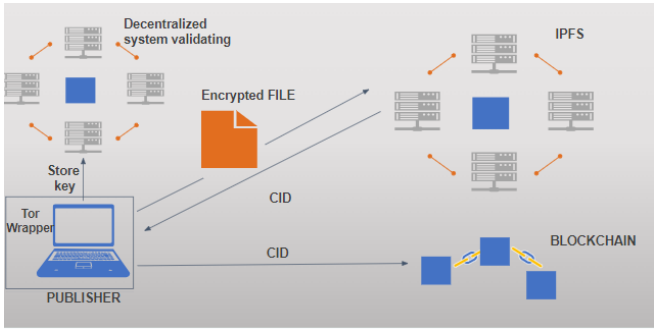

Figure. 3. Sending encrypted file to IPFS network

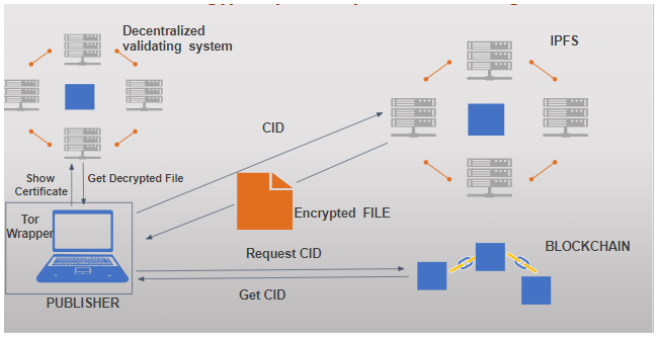

Figure. 4. Accessing encrypted file with CID from IPFS network

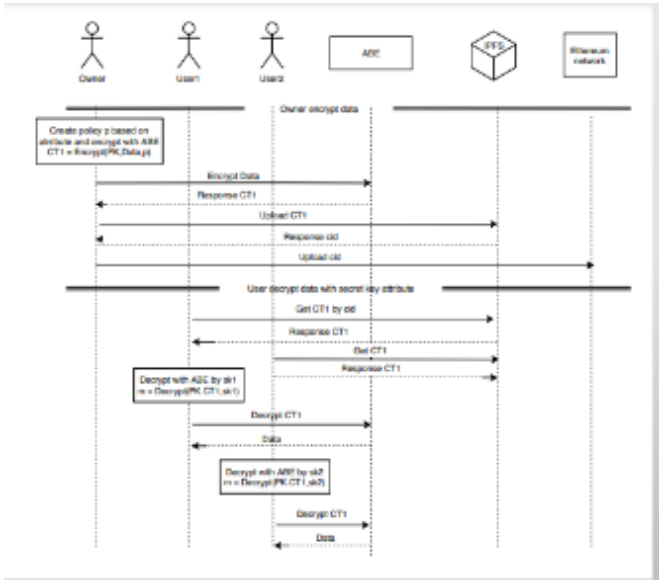

Figure. 5. Attribute based encryption (A.B.E)

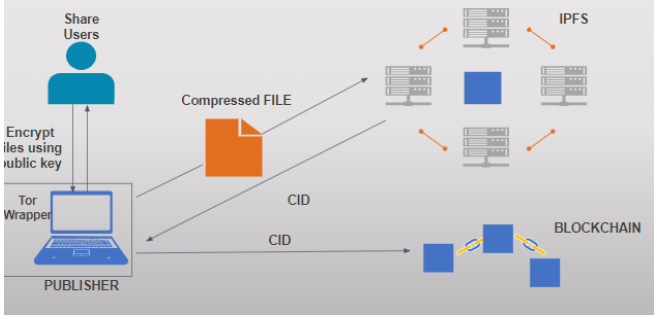

Figure. 6. Sending compressed ZIP file to IPFS network

Technoarete Transactions on Advances in Computer Applications (TTACA)
Vol-2, Issue-3, September 2023
e-ISSN: 2583-3472

RESULTS AND DISCUSSIONS

A. Block Details

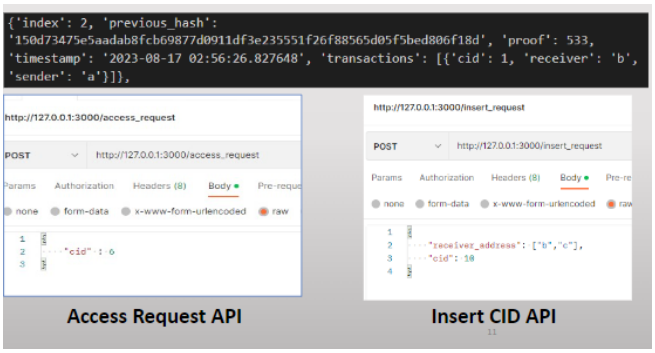

Figure. 7. Access request API with Insert CID API using Postman

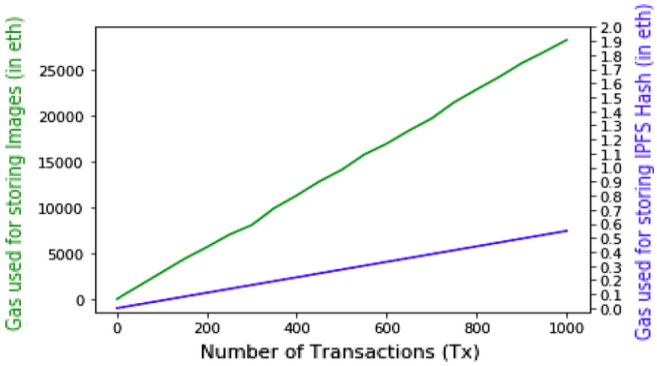

Figure. 8. Graph depicting Gas used on Number of transactions (Tx)

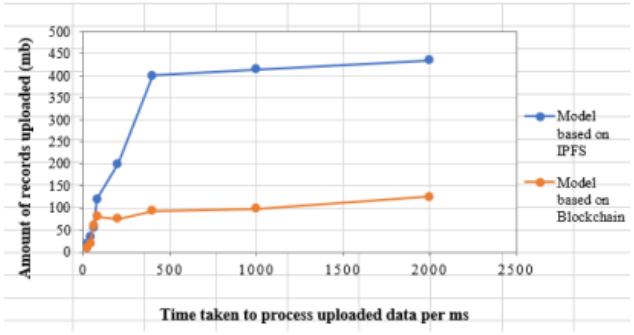

Figure. 9. Amount of files uploaded v/s Time taken to process the data

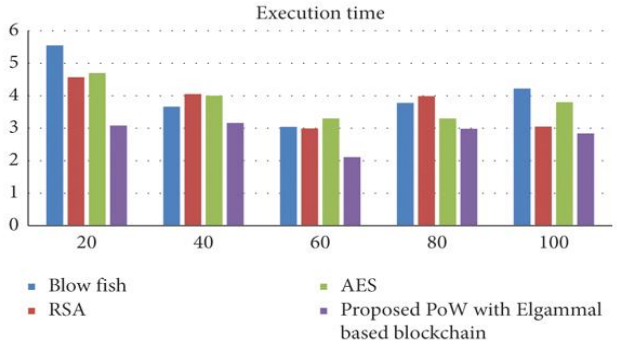

Figure. 10. Execution time of various algorithms like RSA and AES

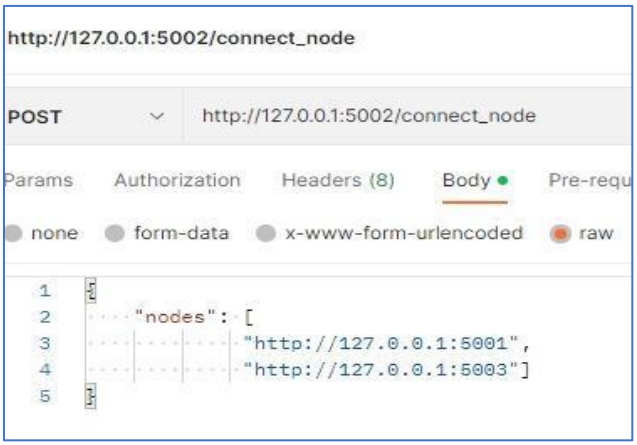

Figure. 11. Node connection with other nodes in the network

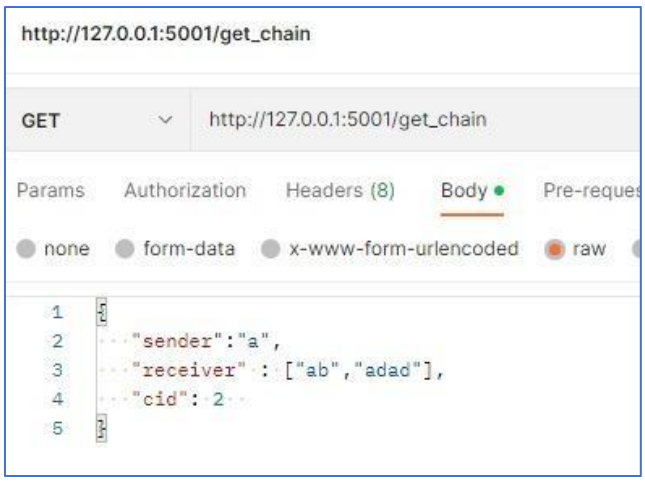

Figure. 12. Details of network having sender, receiver and cid as attributes

CONCLUSION

The need of access control, data protection and anonymity to a user on IPFS can be a primary cause which this projects aims to deliver.In the areas of Journalism this technology (IPFS and Blockchain) can be used to improve the overall security.Improvisations in the limitations of certain areas of blockchain such as cost, data storage and time consumption can take place and this project aims to solve these problems. blockchain technology relies on cryptographic keys to ensure the security and authenticity of transactions. Public and private keys work together to create a secure and transparent ledger, which is maintained by a decentralized network of nodes. The integration of blockchain with IPFS further enhances data security, decentralization, and efficiency by storing data references on the blockchain and the actual data on a distributed and resilient IPFS network. This combination of technologies has the potential to transform the way we store and share digital information, making it more secure, transparent, and accessible to all.

e-ISSN: 2583-3472

FUTURE SCOPE

To have less data redundancy To provide economic incentive to nodes in IPFS by backing it using a crypto currency Working on models to get faster access to files. Supply Chain Management: Blockchain and IPFS can be utilized to enhance transparency and traceability in supply chains. This has the potential to revolutionize industries like food, pharmaceuticals, and electronics, where tracking the origin and journey of products is crucial for safety and authenticity.

Healthcare: Medical records and patient data can be securely stored on a blockchain, ensuring data integrity and enabling patients to have better control over their own health information. IPFS can be used for storing medical images, research data, and other large files efficiently and securely.

Digital Identity: Blockchain technology can be used to create a decentralized and self-sovereign identity system, allowing individuals to have control over their personal information. IPFS can securely store identity-related documents.

REFERENCES

[1] Ashur, T., Dunkelman, O., Talmon, N. (2017). Breaching the Privacy of Israel's Paper Ballot Voting System. In: Electronic Voting. E-Vote-ID 2016. Lecture Notes in Computer Science (), vol 10141. Springer, Cham. https://doi.org/10.1007/978-3-319-52240-1_7.

[2] M. H. Sedky and E. M. Ramzy Hamed, "A secure e-Government's e-voting system," 2015 Science and Information Conference (SAI), London, UK, 2015, pp. 1365-1373, doi: 10.1109/SAI.2015.7237320.

[3] Z. Zheng, S. Xie, H. Dai, X. Chen and H. Wang, "An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends," 2017 IEEE International Congress on Big Data (BigData Congress), Honolulu, HI, USA, 2017, pp. 557-564, doi: 10.1109/BigDataCongress.2017.85.

[4] Nakamoto, S. (2008) Bitcoin: A Peer-to-Peer Electronic Cash System. https://bitcoin.org/bitcoin.pdf

[5] B. Community, "Developer's Guide, Confirmation Score, Transaction Fee and Miner Fee, Minimum Relay Fee, UTXO, Memory Pool, Child Pays for Parent, Raw Transactions," 2018. [Online]. Available: https://bitcoin.org/en/developer-reference#rpc-quick-reference.

[6] Ishaan Anand Srivastava, Bhumika Saini,Dr. Shraddha Phansalkar, Sonali Patwe,"Secure and Transparent Election System for India using Block chain Technology",2019.

[7] V Lalitha, S Samundeswari, R Roobinee, Lakshme S Swetha, "Decentralized Online Voting System using Blockchain", 2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC), INSPEC Accession Number: 21798129, DOI: 10.1109/ICAAIC53929.2022.9792791

[8] Subha P; Padmasree P, Sowndharya Lakshmi R, "Voting System based on BlockChain and using Iris Recognition", 2021 4th International Conference on Computing and Communications Technologies (ICCCT), INSPEC Accession Number: 21722354.

[9] Prof. Mrunal Pathak, Amol Suradkar, Ajinkya Kadam, Akansha Ghodeswar, Prashant Parde, "Blockchain Based E-Voting System", International Journal of Scientific Research in Science and Technology, 2021.

[10] Friðrik Þ. Hjálmarsson, Gunnlaugur K. Hreiðarsson, Mohammad Hamdaqa, Gísli Hjálmtýsson,"Blockchain-Based E-Voting System",2018

[11] M. J. M. Chowdhury, A. Colman, M. A. Kabir, J. Han and P. Sarda, "Blockchain Versus Database: A Critical Analysis," 2018 17th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/ 12th IEEE International Conference on Big Data Science and Engineering (TrustCom/BigDataSE), New York, NY, USA, 2018, pp. 1348-1353, doi: 10.1109/TrustCom/BigDataSE.2018.00186.