

Secure Document storage system using Block chain

**B.Veeru¹, M.Navaneeth², M.Mounika³, A.Vijay⁴, E.Lavanya⁵,
Dr.V .Ramdas⁶**

^{2,3,4,5} B.Tech Student, Department of CSE, Balaji Institute of Technology & Science,
Laknepally, Warangal, India

¹ Assistant Professor, Department of CSE, Balaji Institute of Technology & Science,
Laknepally, Warangal, India

⁶Project Coordinator , Department of CSE, Balaji Institute of Technology & Science,
Laknepally, Warangal, India

Abstract—In contemporary data management systems, ensuring the security and integrity of stored documents is paramount. One prevalent approach leverages blockchain technology, integrating cryptography hashing algorithms such as SHA-256 for enhanced protection. When a document is uploaded, it undergoes hashing to generate a unique cryptography hash. This hash is then embedded into a block chain transaction, alongside relevant metadata, and added to the block chain. Once recorded, the transaction becomes immutable, safeguarding the document's integrity. Verification involves retrieving the hash from the block chain and comparing it with the computed hash of the current document. This process ensures tamper-proof records and authenticates the document's integrity. Utilizing blockchain in this manner establishes a decentralized, secure environment, mitigating risks associated with centralized storage solutions.

I. INTRODUCTION

A reliable and secure document storage solution is more important than ever at a time of digital transformation and the spread of sensitive data. Vulnerabilities in traditional centralised storage solutions include data manipulation, unauthorised access, and single points of failure. Blockchain is a revolutionary technology that has emerged to address these challenges.

Originally designed to serve as the foundation for cryptocurrencies like Bitcoin, blockchain technology has now expanded into a flexible instrument with uses beyond its original intent. Because of its decentralised and unchangeable design, which provides previously unheard-of levels of security, transparency, and trust, it is a viable contender to revolutionise document storage.

At its core, blockchain operates on the principle of distributed ledger technology, wherein a network of decentralized nodes maintains copies of the entire ledger. This distributed nature ensures that no single entity holds control over the entire system, fostering transparency and eliminating the need for a central authority.

In the context of document storage, blockchain-based systems offer several significant advantages. Firstly, the immutability inherent in blockchain ensures that once a document is recorded, it cannot be tampered with or deleted without the consensus of the network. This feature is crucial for maintaining the integrity and authenticity of critical records, legal documents, and sensitive information.

Additionally, the decentralized architecture of blockchain reduces the risk of a single point of failure. Unlike traditional storage systems vulnerable to hacking or system failures, blockchain-based document storage distributes data across a network of nodes, enhancing resilience against such threats.

Even in the event of a compromised node, the integrity of the entire system remains intact, highlighting the robustness and reliability of blockchain technology in securing document storage. As the ecosystem of blockchain-based applications expands, decentralized storage networks like Filecoin play a pivotal role in reshaping the digital landscape. These systems offer a compelling alternative to traditional cloud storage services, empowering users to retain control over their data and reducing reliance on centralized entities prone to censorship, data breaches, and single points of failure. By fostering a decentralized marketplace for storage, Filecoin not only enhances data security but also promotes economic inclusive, enabling individuals worldwide to monetize their unused storage capacity

II. LITERATURE SURVEY

I. Existing Blockchain-Based Document Storage Systems

1. File coin: A Decentralized Storage Network" (2014):

Proposal and Authorship: The concept of Filecoin was first proposed in July 2014 by Juan Benet, the founder of Protocol Labs. The white paper titled "File coin: A Decentralized Storage Network" outlines the foundational principles and objectives of the File coin project. Benet's vision was to create a decentralized storage network that incentivizes participants to contribute storage space in exchange for cryptocurrency rewards.

Description: The white paper introduces File coin as a decentralized storage network built on block chain technology. It elaborates on the inefficiencies of traditional centralized storage solutions and presents File coin as a solution to these challenges. The paper discusses the architecture, consensus mechanisms, and economic incentives of File coin, emphasizing its potential to revolutionize the storage industry. It provides technical specifications and

outlines how File coin addresses security, reliability, and scalability concerns in decentralized storage systems. Overall, the white paper serves as the foundational document that lays out the vision and technical framework for File coin.

2. IPFS - Content Addressed, Versioned, P2P File System" (2015):

Authorship: The paper "IPFS - Content Addressed, Versioned, P2P File System" was authored by Juan Benet, the same individual behind the Filecoin project. Published in 2015, this paper introduces the Inter Planetary File System (IPFS), a protocol designed for decentralized and distributed file storage.

Description: In this paper, Benet presents IPFS as a novel approach to file storage and distribution, addressing the limitations of centralized systems. IPFS utilizes content addressing and peer-to-peer networking to create a distributed file system that is resilient to censorship and single points of failure. The paper outlines the architecture and technical workings of IPFS, including its use of cryptography hashes for content addressing and Merkle DAGs (Directed Acyclic Graphs) for versioning and data integrity. By providing a detailed exploration of IPFS, this paper lays the groundwork for understanding the relationship between IPFS and Filecoin, as both projects share common goals of decentralizing storage infrastructure and improving data accessibility

IBM Blockchain Platform

Overview: IBM provides a comprehensive blockchain platform suitable for a variety of applications, including secure document storage. This platform offers a framework for designing, operating, and governing blockchain networks, enabling organizations to implement tailored solutions for their specific needs. IBM's blockchain platform is known for its robustness, security features, and scalability, making it a popular choice among enterprises seeking to leverage blockchain technology.

II. Challenges and Drawbacks

1. Proof-of-Work Algorithm

One of the common challenges faced by existing block chain systems revolves around the computational expenses associated with proof-of-work algorithms. These algorithms require significant computational resources to calculate random values, impacting the overall efficiency of the system. For example, the Python `random.randint()` function may not offer the desired efficiency required for certain applications.

2. Alternative: Proof-of-Stake Algorithms

To mitigate the computational expenses linked with proof-of-work algorithms, several block chain systems are exploring proof-of-stake algorithms. In these systems, validators are chosen randomly based on the value of the stakes they hold in the block chain. The selected validator gains the right to add a new block to the chain, and the validity of the block affects the value of the validator's stake.

iii. Comparative Analysis

1 Scalability and Transaction Speed

A crucial factor in evaluating block chain-based document storage systems is scalability, particularly concerning transaction speed. As the volume of documents and users increases, systems must efficiently handle transactions to maintain responsiveness. Filecoin, IPFS, and IBM Block chain Platform each offer distinct approaches to addressing scalability.

Filecoin: The decentralized storage network in Filecoin fosters scalability by enabling users to pay for storage services using cryptocurrency. This marketplace model incentivizes providers to expand storage capacity, thereby enhancing system scalability.

IPFS: While not strictly a block chain, IPFS complements block chain technology by providing a decentralized file storage protocol. Its utilization of a distributed hash table contributes to scalability by facilitating efficient content location and sharing among nodes.

IBM Block chain Platform: IBM's platform enhances scalability through its modular architecture, allowing organizations to customize their block chain networks to accommodate increased demands for document storage.

III. EXISTING SYSTEM

As of my last knowledge update in January 2022, the landscape of block chain-based document storage systems and platforms was witnessing significant emergence, with continuous evolution being a defining characteristic. However, it's crucial to acknowledge that developments may have transpired since then, potentially introducing new systems and platforms to the market.

Organizations have been leveraging the IBM Block chain Platform to devise and execute their block chain solutions, thereby augmenting security and transparency in document storage. This platform offers a customization framework tailored to specific organizational requirements, facilitating seamless integration and deployment.

Among the notable examples of existing block chain-based document storage systems as of my last update is Filecoin, renowned for its decentralized storage network. Utilizing blockchain technology, Filecoin establishes a marketplace for storage providers and users, enabling secure and decentralized data storage and retrieval. Users engage storage providers by compensating them in cryptocurrency (Filecoin) for data storage services, with the network employing a combination of blockchain and proof-of-replication mechanisms to ensure data integrity.

Another prominent system is the Interplanetary File System (IPFS), which, while not strictly a blockchain, operates as a decentralized file storage protocol often complementing blockchain technology. IPFS aims to establish a peer-to-peer method for storing and sharing hypermedia in a distributed file system. Leveraging a distributed hash table (DHT), IPFS facilitates content location and sharing among network nodes, thereby reducing reliance on central servers and enhancing decentralization.

Additionally, the IBM Block chain Platform offered by IBM stands out as a versatile block chain framework applicable across various domains, including secure document storage. Organizations can harness this platform to architect and execute block chain solutions tailored to their specific needs, thus reinforcing security, transparency, and efficiency in document storage practices.

IV. PROBLEM STATEMENT

Regarding security, openness, and integrity, current document storage solutions confront significant difficulties. Because they compromise accountability and lack a clear record of changes, centralised models are especially vulnerable. A more reliable access control system and a tamper-proof environment are required due to persistent concerns about unauthorised entry and tampering. Furthermore, constraints in performance and scalability present serious difficulties, particularly when integrating emerging technology. By investigating cutting-edge technologies and developments in the sector, our project aims to revolutionise document storage and provide organisations a safe platform to handle these concerns.

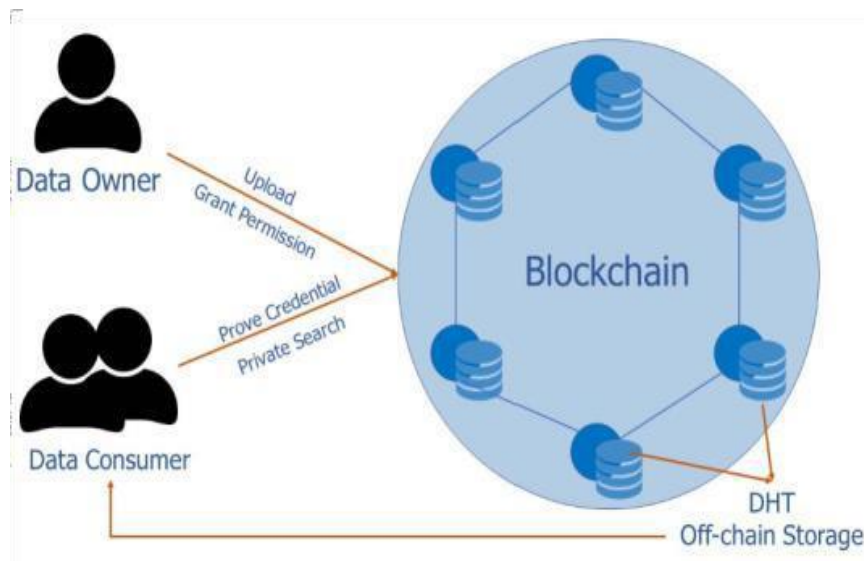


Figure 1: Block chain Architecture

V. PROPOSED METHODOLOGY

With time stamping and encryption, the system ensures data integrity. Its worldwide accessibility enables document verification from any location. Operations may be made more cost-effective and efficient, and they can become more resilient to DDoS assaults. Users also get more privacy control over their cryptographic keys and may provide selective access without disclosing the contents of the full document. Although these benefits are strong, their effective use requires careful consideration of elements like regulatory compliance, scalability, and changing best practices in block chain technology.

ADVANTAGES OF PROPOSED SYSTEM

Peer-to-Peer File Sharing:

Bid adieu to conventional, centralised servers! Our cutting-edge technology removes middlemen and reduces latency for lightning-fast data transmission, enabling direct and smooth file sharing between peers.

Immutable Blockchain Storage:

Embrace our block chain-based file storage solution to embrace the future of data security. Every file that users upload is hashed using cryptography, and safely recorded on the blockchain, establishing an unbreakable chain of transparency and trust.

Decentralised Data Distribution: Take advantage of increased data availability and redundancy as files are automatically shared across many nodes on the block chain network. Goodbye to fears about data loss and hello to a strong distributed file system.

Seamless User Interface: Both tech-savvy users and novices will find our user-friendly interface to be intuitive and seamless. With just a few clicks, share files with ease and take advantage of real-time progress reports when transferring data.

VI. EXPERIMENTAL ANALYSIS

In today's digital landscape, ensuring the security of sensitive documents is paramount. Researchers have recognized the potential of block chain technology in addressing this challenge, prompting experimentation aimed at enhancing document storage security. Here's a detailed analysis:

Objective: The primary objective of these experiments is to evaluate the effectiveness of block chain technology in enhancing document storage security. By leveraging the decentralized and tamper-resistant nature of block chain, researchers aim to provide a robust solution for safeguarding sensitive documents from unauthorized access and modifications.

Methodology: Researchers typically develop and test systems utilizing smart contracts or decentralized protocols specifically designed for secure document storage. These systems often integrate encryption techniques and access control mechanisms to ensure data confidentiality and integrity. The experimentation involves deploying these systems in controlled environments to assess their performance and effectiveness in real-world scenarios.



Fig 2: Flow Diagram

Findings: The experiments consistently demonstrate that block chain offers tamper-resistant storage, significantly reducing the risk of unauthorized access and modifications to stored documents. By leveraging cryptography techniques and decentralized consensus mechanisms, block chain-based document storage systems enhance data security and mitigate the vulnerabilities associated with centralized storage solutions. However, challenges such as scalability and complexity emerge, requiring further exploration and optimization.

Implications: The findings underscore the promising potential of block chain technology in revolutionizing document storage practices. By providing a decentralized, immutable, and transparent storage solution, block chain has the potential to address longstanding security concerns in document management. However, the identified challenges highlight the need for ongoing research and development efforts to fully harness the capabilities of block chain in this domain.

Test Case#	Test Description	Expected Result	Actual Result	Pass/Fail
1	Document Upload	Successfully upload a document	[]	[]
2	Document Encryption	Ensure the uploaded document	[]	[]
3	Blockchain Transaction Initiation	Verify that a blockchain transaction	[]	[]
4	Smart Contract Execution	Confirm that the smart contract manages	[]	[]
5	Document Storage on Blockchain	Validate that the encrypted document	[]	[]
6	Transaction Confirmation	Ensure successful confirmation	[]	[]
7	Document Retrieval	Successfully retrieve a document	[]	[]
8	Access Control Validation	Confirm that access controls enforced	[]	[]
9	Immutability Check	Verify that the stored document	[]	[]
10	Transparent and Auditable Record	Ensure blockchain provides transparent	[]	[]
11	Decentralization and Resilience	Test resilience by attempting retrieval	[]	[]
12	Compatibility with Existing Systems	Confirm seamless integration	[]	[]
13	Scalability Test	Assess the system's ability	[]	[]
14	Traceability and Accountability	Validate the traceability of document	[]	[]
15	Prototype Testing	Evaluate the effectiveness	[]	[]

Figure 3; Experimental Analysis

Conclusion: The experiment analysis highlights blockchain's significant potential for secure document storage. While the findings showcase the inherent strengths of blockchain in enhancing data security, they also underscore the importance of addressing scalability and complexity challenges. Continued innovation and collaboration across academia, industry, and government sectors are essential to unlock the full benefits of blockchain technology in document storage and management.

VI. CONCLUSION

In conclusion, the creation and use of a blockchain-based secure document storage system marks a substantial advancement in resolving the shortcomings and restrictions of conventional storage systems. The use of blockchain technology provides exceptional security, transparency, and decentralisation, making it a dependable option for protecting confidential data. The main elements of the suggested system—smart contracts, decentralised storage, and access controls—cooperate to provide a complete and safe environment for document archiving. Using smart contracts and cryptographic keys, access control techniques provide a granular degree of security that enables organisations to set and enforce stringent guidelines for document access and change. This improves system security overall by guaranteeing that only authorised parties may interact with sensitive information. Essentially, a blockchain-based secure document storage solution not only solves today's data security issues but also paves the way for an approach to information management that is more transparent, decentralised, and resilient. With businesses realising the value of protecting their most important assets, data, blockchain-based solutions provide a transformative route to a more reliable and secure digital environment.

VII. REFERENCES

1. Ramdas Vankdothu, Dr.Mohd Abdul Hameed “A Security Applicable with Deep Learning Algorithm for Big Data Analysis”, *Test Engineering & Management Journal*, January-February 2020
2. Ramdas Vankdothu, G. Shyama Chandra Prasad “ A Study on Privacy Applicable Deep Learning Schemes for Big Data” *Complexity International Journal*, Volume 23, Issue 2, July-August 2019
3. Ramdas Vankdothu, Dr.Mohd Abdul Hameed, Husnah Fatima “ Brain Image Recognition using Internet of Medical Things based Support Value based Adaptive Deep Neural Network” *The International journal of analytical and experimental modal analysis*, Volume XII, Issue IV, April/2020
4. Ramdas Vankdothu, Dr.Mohd Abdul Hameed, Husnah Fatima” Adaptive Features Selection and EDNN based Brain Image Recognition In Internet Of Medical Things “ *Journal of Engineering Sciences*, Vol 11, Issue 4 , April/2020 (UGC Care Journal)
5. Ramdas Vankdothu, Dr.Mohd Abdul Hameed “ Implementation of a Privacy based Deep Learning Algorithm for Big Data Analytics”, *Complexity International Journal* , Volume 24, Issue 01, Jan 2020
6. Ramdas Vankdothu, G. Shyama Chandra Prasad” A Survey On Big Data Analytics: Challenges, Open Research Issues and Tools” *International Journal For Innovative Engineering and Management Research*, Vol 08 Issue 08, Aug 2019
7. Ramdas Vankdothu, Dr.Mohd Abdul Hameed, Husnah Fatima” A Brain Tumor Identification and Classification Using Deep Learning based on CNN-LSTM Method” *Computers and Electrical Engineering* , 101 (2022) 107960
8. Ramdas Vankdothu, Mohd Abdul Hameed “Adaptive features selection and EDNN based brain image recognition on the internet of medical things”, *Computers and Electrical Engineering* , 103 (2022) 108338.

9. Ramdas Vankdothu, Mohd Abdul Hameed, Ayesha Ameen, Raheem, Unnisa “ Brain image identification and classification on Internet of Medical Things in healthcare system using support value based deep neural network” Computers and Electrical Engineering, 102(2022) 108196.
10. Ramdas Vankdothu, Mohd Abdul Hameed” Brain tumor segmentation of MR images using SVM and fuzzy classifier in machine learning” Measurement: Sensors Journal, Volume 24, 2022, 100440
11. Ramdas Vankdothu, Mohd Abdul Hameed” Brain tumor MRI images identification and classification based on the recurrent convolutional neural network” Measurement: Sensors Journal, Volume 24, 2022, 100412 .
12. Bhukya Madhu, M.Venu Gopala Chari, Ramdas Vankdothu, Arun Kumar Silivery, Veerender Aerranagula ” Intrusion detection models for IOT networks via deep learning approaches ” Measurement: Sensors Journal, Volume 25, 2022, 10064
13. Mohd Thousif Ahemad , Mohd Abdul Hameed, Ramdas Vankdothu” COVID-19 detection and classification for machine learning methods using human genomic data” Measurement: Sensors Journal, Volume 24, 2022, 100537
14. S. Rakesh ^a, Nagaratna P. Hegde ^b, M. Venu Gopalachari ^c, D. Jayaram ^c, Bhukya Madhu ^d, Mohd Abdul Hameed ^a, Ramdas Vankdothu ^e, L.K. Suresh Kumar “Moving object detection using modified GMM based background subtraction” Measurement: Sensors Journal, Volume 30, 2023, 100898
15. Ramdas Vankdothu, Dr. Mohd Abdul Hameed, Husnah Fatima “Efficient Detection of Brain Tumor Using Unsupervised Modified Deep Belief Network in Big Data” Journal of Adv Research in Dynamical & Control Systems, Vol. 12, 2020.
16. Ramdas Vankdothu, Dr. Mohd Abdul Hameed, Husnah Fatima “Internet of Medical Things of Brain Image Recognition Algorithm and High Performance Computing by Convolutional Neural Network” International Journal of Advanced Science and Technology, Vol. 29, No. 6, (2020), pp. 2875 – 2881
17. Ramdas Vankdothu, Dr. Mohd Abdul Hameed, Husnah Fatima “Convolutional Neural Network-Based Brain Image Recognition Algorithm And High-Performance Computing”, Journal Of Critical Reviews, Vol 7, Issue 08, 2020
18. Antonopoulos, A.M. (2014). Mastering Bitcoin: Unlocking Digital Cryptocurrencies. O'Reilly Media.
19. Drescher, D. (2017). Blockchain Basics: A Non-Technical Introduction in 25 Steps. Après.
20. Tapscott, D., & Tapscott, A. (2016). Blockchain Revolution: How the Technology Behind Bitcoin and Other Cryptocurrencies is Changing the World. Penguin
21. Swan, M. (2015). Blockchain: blueprint for a new economy. O'Reilly Media, Inc.
Narayanan, A., Bonneau, J., Felten, E., Miller, A., & Goldfeder, S. (2016). Bitcoin



I am M. Navaneeth from Department of Computer Science and Engineering. Currently, pursuing final year at Balaji Institute of Technology and Science. My research is done based on “SECURE DOCUMENT STORAGE SYSTEM USING BLOCKCHAIN”.



I am M. Mounika from Department of Computer Science and Engineering. Currently, pursuing final year at Balaji Institute of Technology and Science .My research is done based on “SECURE DOCUMENT STORAGE SYSTEM USING BLOCKCHAIN”.



I am A. Vijay from Department of Computer Science and Engineering. Currently, pursuing final year at Balaji Institute of Technology and Science .My research is done based on “SECURE DOCUMENT STORAGE SYSTEM USING BLOCKCHAIN”.



I am E. Lavanya from Department of Computer Science and Engineering. Currently, pursuing final year at Balaji Institute of Technology and Science .My research is done based on “SECURE DOCUMENT STORAGE SYSTEM USING BLOCKCHAIN”.

I