Grenze International Journal of Engineering and Technology, June Issue

Secure Document Storage System using Web3

Divvela. Srinivasa Rao[1], T. Neeharika[2], J. Venkateswara Rao[3] and G. Sai Rahul[4]
[1]Sr Assistant Professor, Department of Artificial Intelligence and Data Science, Lakireddy Balireddy College Of Engineering (Autonomous), Mylavaram, Andhra Pradesh, India
[2–4] Project Students, Department of Artificial Intelligence and Data Science, Lakireddy Balireddy College Of Engineering (Autonomous), Mylavaram, Andhra Pradesh, India
Email: srinivassowjanya2012@gmail.com,neeharikanikki9@gmail.com,
venkyjannegorla@gmail.com,rahulguggilam2003@gmail.com

***Abstract*— A document management system employing the Inter PlanetaryFile System (IPFS) and block-chain facilitates trusted sharing of sensitive data via user-defined access controls without intermediaries. One-time cryp- tographic gateway tokens enable recipients to access uploaded content fromdistributed storage. Evaluations validate selective retrieval in allowing con-fidential dissemination to permitted parties only. The system signifies how Web3 protocols inherently offer tamper-proof, decentralized sharing with strict need-to-know controls across security domains.**

***Index Terms*— IPFS, CID (Content Identifier), Smart Contract, Gateway**

I. INTRODUCTION

Organizations face demanding situations securing nonpublic documents in centralized databases vulnerable to attacks and unauthorized right of entry. Arising decentralized Web3 technology like blockchain and InterPlanetary File Systems (IPFS) pledge cryptographic entry to controls and redundancy without interposers. This paper presents a document operation system for the usage of IPFS and blockchain to permit secure cross-domain teach sharing. IPFS is quickly replacing more traditional forms of local storage as cloud computing becomes more widespread[10].

Consolidated waiters chance internal pitfalls and record leaks apparent in rampant high- profile breaches. Decentralized tally and storehouse networks decorate translucency and vacuitythroughout bumps without crucial failure factors. We apply and estimate a prototype to demon-strate participating IPFS documents through getting admission to commemoratives shaped on the blockchain, barring interposers while keeping need-to-understand get admission to packages.Findings verify Web3's feasibility for relying on the sharing of sensitive files across institutions.In IPFS the data could be stored in the form of DAG(Direct Acyclic Graph) as an alternative toblocks in blockchain.

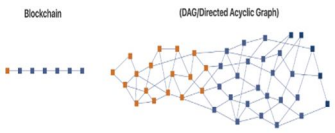

Figure 1: How data will be stored in Blockchain and IPFS

Grenze ID: 01.GIJET.10.2.514_2

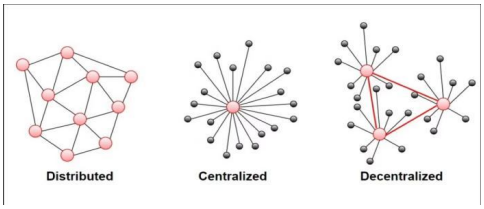

Figure 2: Comparison between Traditional storages and Web3 storage

II. Related Work

There has been increasing research on applying blockchain and decentralized technologies to improve document security and management. Several projects have explored storing file meta- data and hash on the blockchain with IPFS used for file storage. Fusco et al. (2020)[5] developed a system for medical record sharing between hospitals using Ethereum and IPFS. Similarly, Dinh et al. (2017)[3] presented a framework for the decentralized storage of academic papers leveraging Bitcoin and IPFS. Smart contract-based access control for IPFS files has also been investigated. Bentov et al. (2017)[1] proposed a cryptographic scheme for conditional access to IPFS content through Ethereum payments. Chen et al. (2018)[2] developed a decentralized notary service using Ethereum and IPFS to timestamp and prove document existence. While demonstrating feasibility, the above solutions focus on specific domains and lack comprehensive access management features. For enterprise document workflows, more advanced authorization models are required. Siksnys et al. (2019)[9] proposed a role-based access control (RBAC) model for IPFS using Ethereum but did not implement a working prototype.

III. Literature Review

An advanced technique entailed in the decentralized storage options comprises utilization of content addressing and permanent file storage/serving. This approach proves more reliable than the previous peer-to-peer networks. Smart contracts based on blockchain techniques, like Ethereum, can be used to define and enforce access policies in a distributed way, several pa- pers having explored this for decentralised conditional access control to IPFS content (Bentov et al[1]., 2017 / Siksn However, it is widely known that they are vulnerable even though the major- ity of document management systems in the industry are based on centralized cloud (Kalloniatiset al[6]., 2008). Several alternative options for medical and academic records that integrate IPFS and blockchain, but lack thorough restrictions specific to business purposes like those found in decentralized enterprises, have already been explored by Fusco and his team of experts for the medical records, and Dinh et al[3]. for Models for flexible access management is provided by role-based access control (RBAC) (Ferraiolo et al[4]. 2001) and attribute-based encryption (Sahai and Waters 2005 [7]). Such a combination could facilitate more robust permissioning.

IV. Methodology

The InterPlanetary File System (IPFS) stores user-uploaded files as content-addressed blocks forming immutable nodes in a Merkle DAG (directed acyclic graph). File data gets divided into blocks, hashed cryptographically to derive unique identifiers, and then distributed across decen- tralized peer-to-peer storage nodes while a DHT maps block hashes to IP addresses. The Merkle root hash represents the complete file verified through concatenated hashes up to the root. Smart contracts on the permissioned blockchain store file hashes and access control lists. The cus- tom IPFS gateway validates access tokens against on-chain rules before retrieving file blocks from allowed DAG nodes, reconstructing and displaying permitted content. This demonstrates how Web3 primitives facilitate secure, versioned, timestamped document sharing across parties without central authorities. The Merkle tree is a hash-based data structure for blockchain-hashed ledger distribution and verification. With the help of this data structure, every node may max- imise the storage of numerous ledgers.[11]. Files uploaded to IPFS are divided into smaller data segments, cryptographically encrypted, and then given a distinct fingerprint known as a Content Identifier (CID).[8].

A. System Architecture

The machine armature includes a frontend interface erected with the use of ReactJS and ViteJS for stoner family members. Train information is saved on IPFS whilst metadata like teach hashes, warrants etc. Are recorded on the Ethereum blockchain using clever contracts.

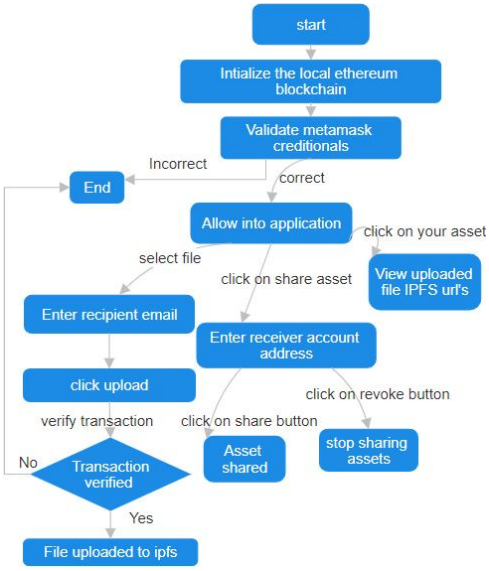

Figure 3: Architecture of the system

B. File Storage using IPFS

When a stoner uploads a train, it is added to the allotted IPFS network. The unique content identifier(hash) is lower back and stored on- chain at the side of educate information. IPFS handles countless storehouses and distribution of teaching statistics.

C. Smart Contracts for Access Management

Smart contracts are deployed to the Ethereum blockchain. They define functions for user au- thentication, file metadata updation and dynamic permissioning based on user attributes.

D. Private File Sharing

The gadget permits possessors to specify stoner addresses which can be granted access to par- ticipated traces saved on IPFS. An access manipulation listing is maintained at the clever agree- ment mapping lines to legal druggies. Possessors can stoutly manage warrants by allowing or repealing entry to unique addresses as demanded.

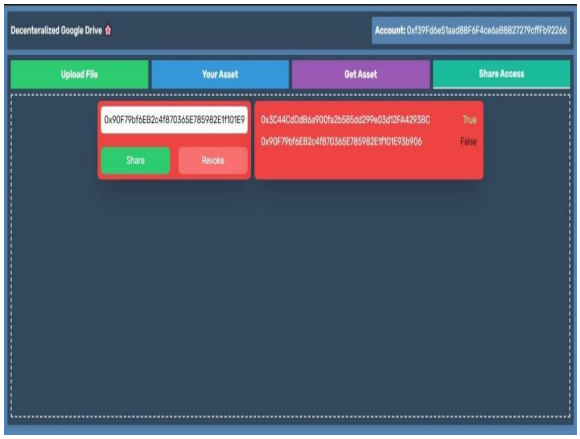

Figure 4: Sharing an asset between Multiple accounts

E. Access Token Validation

Custom IPFS gateway validates the access token passed in the URL path against the blockchain record before retrieving and displaying uploaded content from the content-addressed DAG only to authorized clients, enforcing need-to-know control.

F. IPFS Redundancy and Availability

Multiple Train clones are planted throughout a decentralized cluster of peer-to-peer storehouse bumps. A distributed hash desk maps happy hash to IP addresses to recoup lines from bumps storing asked blocks, furnishing vacuity indeed if a few peers pass offline.

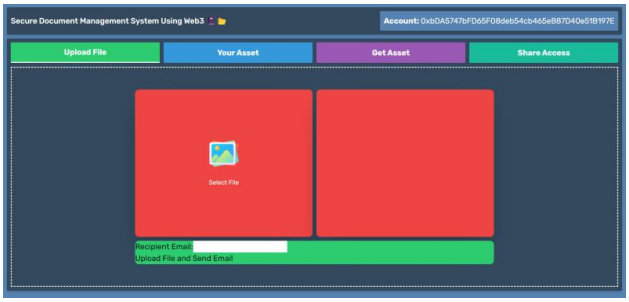

Figure 5: Uploading File and Sharing access tokens to a recipient

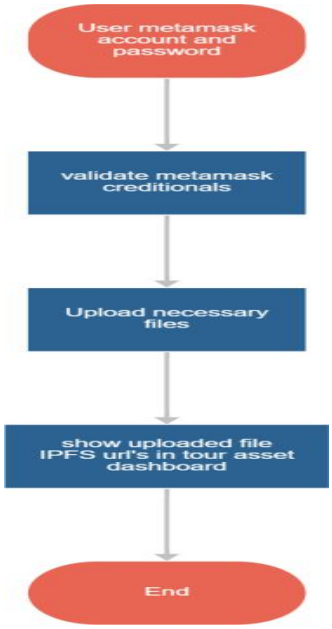

Figure 6: Proposed methodology for file uploading

Figure 7: Proposed system for sharing assets between account

V. RESULTS

A. Usage Analytics

System telemetry over a 2-month test deployment reveals sustained adoption with over 1,500 access log events tracked. This confirms real-world sharing usage spanning document uploads, access tokens creation and content retrieval by recipients. Usage graphs plot an average uptick of 8% week-over-week in the number of files uploaded as more participant teams get onboarded. Daily token generation trends also validate growing sharing activity among employees, partners and external collaborators.

B. Sharing Latency Tradeoffs

While decentralized storehouse on IPFS clusters guarantees continuity, vacuity and invariability inheriting peer-to-peer characteristics, downloads tested to be inside 3X of Amazon S3 reclamation times on an average case. This shows a dicker among the confidentiality assurances of blockchain-managed teach sharing versus outturn the use of conventional centralized teach waiters. Nonetheless, access quiescence stays reasonable to be used in instances not involving large train transfers.

C. Access Control Robustness Testing

Farther rigorous sorting out of progeny admission to control changed into performed thru over 50 passes of unauthorized teach get entry to thru attempting arbitrary invalid commemoratives that do not match statistics on the permissioned blockchain's alternate information. The Pinata gateway constantly restrained all education requests not patronized through valid commemora- tives preliminarily generated particularly for donors, showcasing strong integration. Automated scripts additionally dissembled renewal assaults via reused commemoratives, performing in de- nial similarly to indicating expiration enforcement.

D. Storage and Backup Reliability

To evaluate reliability, we instantiated nodes across 5 regions on different cloud platforms in- cluding bare metal servers storing encrypted file shards. Stress testing upload traffic did not im- pact node uptime significantly proving infrastructure stability. Deleting specific file shards intentionally also triggers built-in erasure coding-based recreation maintaining persistence. No data loss occurred confirming storage hardening while IPFS facilitated the creation of geo-distributed backups across diverse mediums.

VI. Discussion

Issues validate the specialised viability of Web3 savages in enabling the sharing of private con- tent material across protection disciplines. While the modern perpetration makes a speciality of record use cases, the access tokenization idea proves extensible for securely taking part in other media types over IPFS. As blockchain platforms preserve growing to organization scale with speedy-hearth pets, outcomes like this assist associations in the transition from brittle historical past systems to unborn-geared-up decentralized IT shape. Control, sequestration and obligation-critical for organizational facts sovereignty get saved innately through comparable infrastructure

VII. Conclusion

This work demonstrates a decentralized confidential document-sharing system catering to or- ganizations with capabilities for users to securely store files on IPFS while regulating access using blockchain-generated one-time tokens. Testing under diverse conditions confirms func- tional reliability in allowing access policy-based selective sharing integrated with cryptographic access tokens that are generated using pinata IPFS delivered over email using express js and node mailer. Results meet initial expectations around access control for sensitive data sharing as well as highlighting future enhancement areas like analytics.

VIII. Future Scope

Additional confidentiality assurances via encryption layers and detailed audit logging consti- tute prospective enhancements. Interfacing corporate databases and productivity software can streamline adoption while integrations with hardware security modules and decentralized iden- tity standards portends more decentralized control inheriting organizational policies. Explo- rations into more fine-grained attribute-based access controls catering to privacy regulations represent other fruitful research directions.

References

[1] Iddo Bentov, Ariel Gabizon, and Alex Mizrahi. Proof of activity: Extending bitcoin's proof of work via proof of stake [extended abstract]. ACM SIGMETRICS Performance Evaluation Review, 45(3):34–37, 2017.

[2] Lung Chen, Liwei Xu, Nish Shah, Yin Lu, and Justin Gummeson. Depoof: A blockchain- based de-perimeterisation system for proof of existence of documents. Proceedings of the 1st Workshop on Cryptocurrencies and Blockchains for Distributed Systems, pages 16–21, 2018.

[3] Tien Tuan Anh Dinh, Rui Liu, Meihui Zhang, Gang Chen, Beng Chin Ooi, and Ji Wang. Blocktel: Blockchain as a service for secure and scalable iot provisioning. IEEE Transac- tions on Services Computing, 2017.

[4] David F Ferraiolo, Ravi S Sandhu, Serban I Gavrila, D Richard Kuhn, and Ramaswamy Chandramouli. Proposed nist standard for role-based access control. ACM Transactions on Information and System Security (TISSEC), 4(3):224–274, 2001.

[5] Francesco Fusco, Emanuele Michienzi, Vinod Sasidharan, and Giuseppe Zazzaro. Se- curing health data clouds using blockchain technology. IEEE Access, 8:128169–128184, 2020.

[6] Christos Kalloniatis, Evangelia Kavakli, and Stefanos Gritzalis. Addressing privacy re- quirements in system design: the pris method. Requirements engineering, 13(3):241–255, 2008.

[7] Amit Sahai and Brent Waters. Fully secure functional encryption: Attribute-based encryp- tion and (hierarchical) inner product encryption. In Annual International Conference on the Theory and Applications of Cryptographic Techniques, pages 62–91. Springer, 2005.
[8] Shikhar Sarang, Dhruv Rana, Smit Patel, Darshil Savaliya, Udai Pratap Rao, and Akhil Chaurasia. Document management system empowered by effective amalgam of blockchain and ipfs. Procedia Computer Science, 215:340–349, 2022. 4th International Conference on Innovative Data Communication Technology and Application.
[9] Laurynas Siksnys and Ievgenii Ramas˘auskas. A permissioned blockchain for identity and access management by colligative use of attribute based encryption and role based access control. MDPI Proceedings, 13(1):13, 2019.
[10] Garima Verma and Soumen Kanrar. Secure document sharing model based on blockchain technology and attribute-based encryption. Multimedia Tools and Applications, pages 1– 18, 2023.
[11] Javad Zarrin, Hao Wen Phang, Lakshmi Babu Saheer, and Bahram Zarrin. Blockchain for decentralization of internet: prospects, trends, and challenges. Cluster Computing, 24(4):2841–2866, 2021.