Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.1, January 2024, pp. 02401017-01 ~ 02401017-07
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i1.17772

Developing Decentralized Data Storage Network Using Blockchain Technology to Prevent Data Alteration

Ryan Adi Putra*, Rizka Ardiansyah, M. Yazdi Pusadan, Anita Ahmad Kasim, Yuri Yudhaswana Joefrie

Information Technology Department, Faculty of Engineering, Tadulako University, Jl. Soekarno Hatta No KM 9 Palu 94148, Central Sulawesi, Indonesia

*ryannadiputraa@gmail.com

Abstract. In the face of escalating global data exchange, the pronounced vulnerability of traditional centralized storage networks to manipulation and attacks poses a pressing challenge. Digital service providers, entrusted with vast datasets, grapple with the formidable task of ensuring the security, integrity, and continuous availability of their stored information. This paper tackles these multifaceted issues by proposing a decentralized data storage network empowered by blockchain technology. This approach systematically mitigates the inherent susceptibilities of centralized systems, thereby providing heightened resilience against unauthorized alterations and malicious attacks that compromise digital information integrity. Moreover, the decentralized model holds significant promise for securing public data. By leveraging the transparency and immutability of blockchain ledgers, this approach not only safeguards against unauthorized access but also actively fosters transparency and accountability in data management. This makes it particularly well-suited for ensuring the security and integrity of public data, addressing concerns related to trust and reliability in the ever-evolving landscape of information exchange.

Keywords: Blockchain, Decentralization, Data Storage, Data Security

(Received 2023-12-14, Accepted 2023-12-24, Available Online by 2024-01-03)

1. Introduction

The exchange of data has experienced a significant global increase compared to the last few decades. The management of data in the digital realm has become an integral part of various industrial sectors. The rise in volume and complexity of data generated by various digital service providers poses challenges in ensuring the security, integrity, and availability of data. The conventional approach of storing data in centralized networks poses significant obstacles in terms of security, integrity, and accessibility. One critical challenge that has garnered attention is the vulnerability of centralized systems to data manipulation and alteration. The inherent centralization makes these systems susceptible to unauthorized access and malicious activities, leading to potential compromises in data integrity. Decentralized platform systematically mitigates these issues by encrypts and distributes data across a decentralized network, with the use of blockchain technology as a distributed ledger to verify and

authenticated every transcation in the network using public-key encryption and consensus protocol [1-3].

Blockchain is a technology that amalgamates various disciplines, including computer science, mathematics, and cryptography. Blockchain technology enables the formation of a decentralized data storage network resistant to changes, manipulation, or network-based attacks [4-6]. Blockchain technology is characterized by trustlessness, eliminating dependence on centralized storage providers through the implementation of a transparent and decentralized system. The decentralized system does not involve third parties or intermediaries; instead, it utilizes a peer-to-peer (P2P) method allowing direct communication between nodes, and with the consensus protocol that eliminates the potential for data fraud in the network [7], [8].

This research aims to develop a decentralized data storage network infrastructure by implementing blockchain technology to prevent data manipulation by centralized data storage. The findings of this research are expected to contribute insights into the development of blockchain as a decentralized data storage solution.

2. Methods

2.1. Type of Research

This research employs an exploratory research type and qualitative method. Exploratory research examines existing knowledge on a particular topic and other relevant information. The aim of this research is to provide fundamental knowledge on a topic as an introduction for further studies. And by employing qualitative methods, this research seeks to offer in-depth explorations that serve as an insightful introduction, laying the groundwork for subsequent studies [9].

2.2. System Development

The system development is carried out using the prototype method. The programming language used is Go for the blockchain nodes, and Typescript with the NestJS framework for the API Gateway. Badger DB is utilized as persistent and fast key-value database on each node in the blockchain network. Inter-node communication in the blockchain network is facilitated through the gRPC protocol, while REST API is employed in the API Gateway. Containerization is achieved using Docker to simplify system deployment and configuration [10]. The use of the prototype method allows developers to design a running prototype as a sample of the developed system.

3. Results and Discussion

3.1. Planning

The planning phase begins with identifying issues and conducting a literature review on both centralized data storage systems and decentralized systems as a solution to the identified problems. The exploratory research method is then determined by implementing fundamental aspects of decentralized blockchain systems based on related studies.

3.2. Analysis

During this phase, it was found that the decentralized nature not only mitigates the risk of a single point of failure but also ensures data immutability through cryptographic mechanisms. Additionally, the increased transparency and trust derived from blockchain technology contribute significantly to the reliability of data transactions. Despite the promising benefits, the analysis reveals noteworthy challenges in the development and maintenance of decentralized data storage with blockchain. The intricate nature of blockchain technology demands specialized expertise, making system development more resource-intensive. Interoperability issues and scalability concerns also emerge as potential challenges, requiring careful consideration to ensure seamless integration and sustained performance.

3.3. System Design

The system network infrastructure can be seen in the Figure 1, consists of multiple nodes to handle blockchain transcations built based on microservice architecture, with each node run within an internal network. These nodes can later be configured and added to scale the network [11]. Inter-node communication is facilitated using the gRPC protocol, which is an RPC (Remote Procedure Call) protocol developed by Google. gRPC utilizes Protocol Buffers (protobuf) as the Interface Definition Language (IDL), and it operates over HTTP/2, making inter-node communication faster and more efficient [12]. Clients make HTTP calls to an API Gateway, which then forwards the requests to be processed by the blockchain network through NGINX as a load balancer [13], [14].

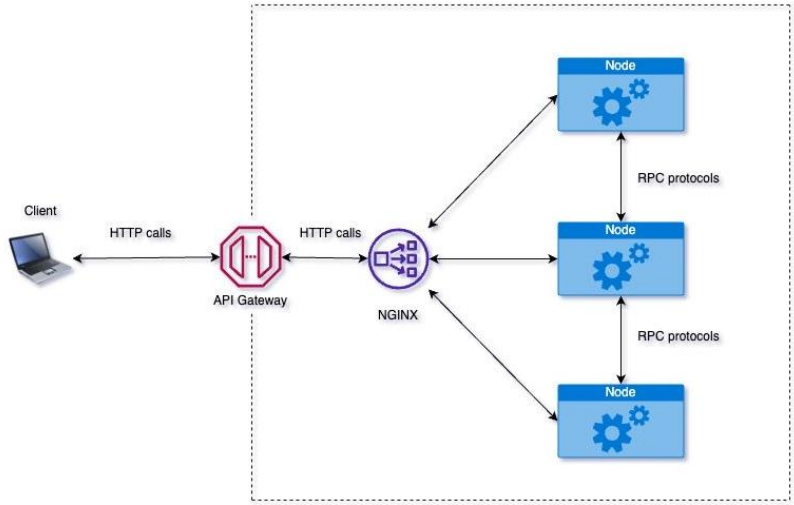

Figure 1. Blockchain Data Storage Network System

3.4. Prototype

The prototype utilizes a blockchain node capable of receiving, hashing, and storing data in JSON format. This prototype features the primary function of blockchain storage in the system under development.

3.5. Implementation

The system development is divided into two components: the blockchain node where data is stored as blocks in the blockchain network, and the API Gateway that connects applications to the blockchain network. The hashing algorithm used in blockchain storage employs the SHA256 algorithm with the Proof of Work (POW) consensus algorithm. In the Proof of Work (POW) consensus algorithm, nodes in the network must use a significant amount of computing resources to add a block to the blockchain. This computation involves a hashing process that combines the hash of the previous data, current data and a random value. As shown in Figure 2 the new block depends on the previous block hash, this hashing process ensures that any changes to the data will result in a different hash outcome [15-18].

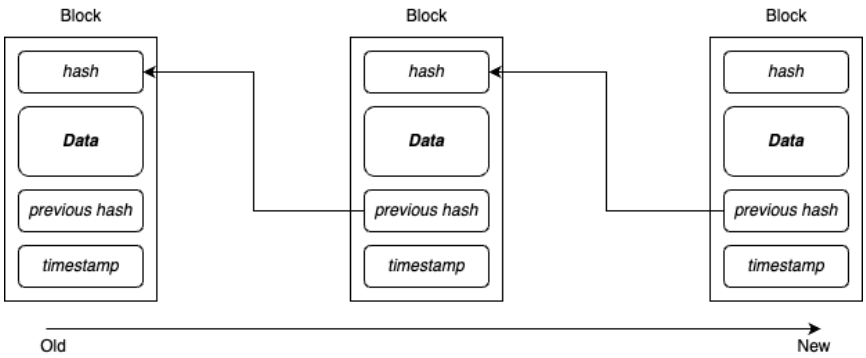

Figure 2. Block Data Structure

Each node in the blockchain network receives data in JSON format from NGINX as a load balancer. The node receiving the data executes the consensus algorithm and adds the data as a block in the blockchain, then sends the data to other nodes using gRPC to be stored in their respective blockchains. The communication mechanism employed is the Gossip Protocol, allowing each node in the blockchain network to receive data received from the API, additionally the system will check the previous hash wether it matchs or not [19], [20]. If previoush hash didn't match with the hash that stored in the node, it will then retrieve the full copy of blockchain from other nodes to repair itself, this aims to prevent data manipulation that purposely change the data block in the blockchain. During the development phase, docker compose were utlize to run multiple service in the blockchain network inside Docker container, while also simulate the full function of the system networks, the example of configurations can be seen on Figure 3.

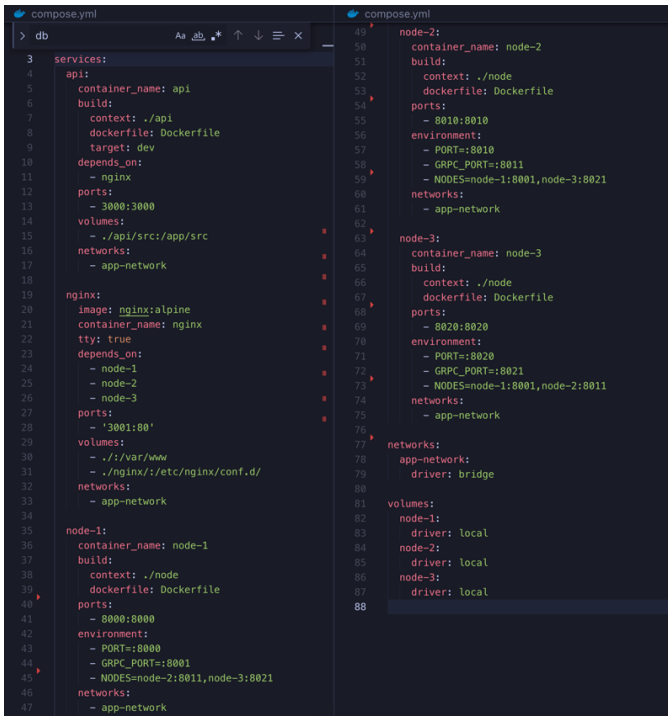

Figure 3. Docker Compose Configuration

The creation of the API Gateway serves as a bridge between applications and the blockchain network as a decentralized data storage, using NodeJS with the NestJS framework. The API Gateway forwards

HTTP requests to NGINX as a load balancer, which is subsequently processed by nodes in the blockchain network. There are three endpoints for data transactions in the blockchain network through the API Gateway, including adding and retrieving data from blocks in the blockchain network, as well as retrieving the full copy of blockchain. Data payload can be any type as long it is a valid JSON object.

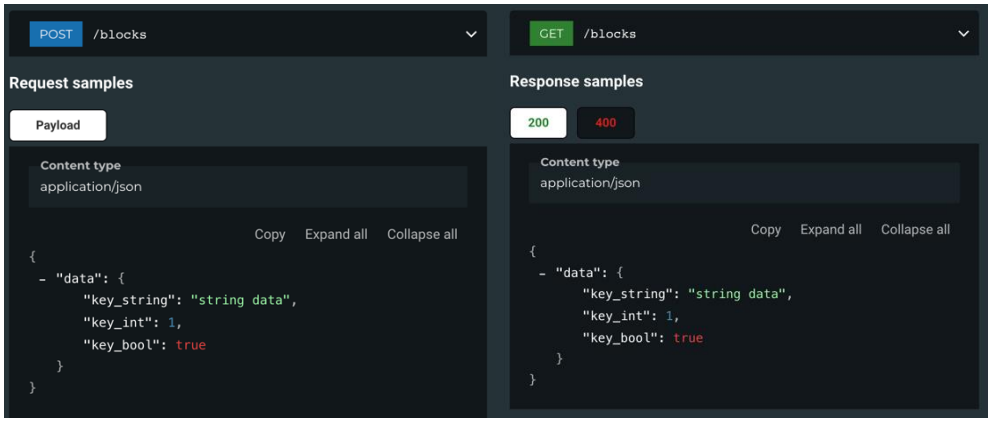

Figure 4. Send Block and Get Latest Block Request

Figure 4 and 5 show the example of API Request and Response on the API Gateway, the API will accept any kind of data structure as long as it was sends in JSON format, this will enable client to be flexible for storing their data. The "/blocks" endpoint is utilized to add data using the POST method, meanwhile, the GET method returns the latest block from the blockchain network. Additionally, the "/blockchains" endpoint returns the full copy of blockchain, containing the history from the first or genesis block to the latest block in the blockchain.

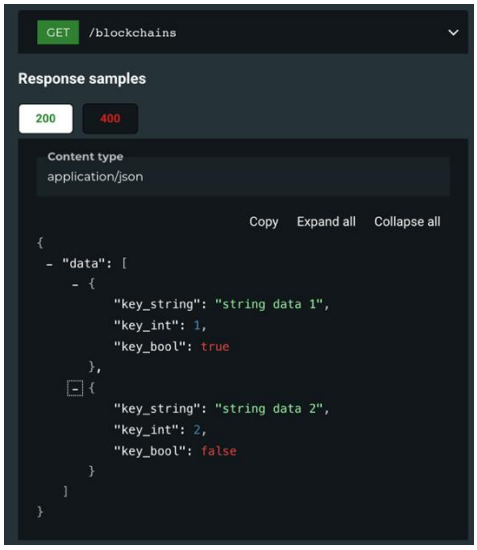

Figure 5. Retrieve The Full Copy of Blockchain

3.6. System Testing

Testing conducted using black box testing method to ensure the developed system meet the specified decentralized data storage system requirements. The system will be tested by matching the expected output and the result by each API endpoints. Testing will also include the data manipulation simulation in a blockchain node, to verified the system capability to endure data manipulation.

Table 1. System Testing Results

Test Case	Description	Result
Add data to system	Send HTTP POST request with JSON data payload that will then addeds data in blockchain network	Success
Retrieve latest data	Send HTTP GET request that will return latest data from blockchain network	Success
Retrieve full copy of blockchain	Send HTTP GET request that will return full copy of blockchain that contain data from genesis block to the latest	Success
Alter data from blockchain	Manually modify the data from the blockchain and then run the system to allow nodes to repair themselves when consensus fails during a transaction by downloading a valid copy of the blockchain from the nearest configured blockchain node	Success
Alter data from the middle of blockchain nodes (change data history)	Manually change data from the middle block on the blockchain then run the system to allow node to repair itself when consensus is failing since the hash and previous hash block didn't match	Success

4. Conclusion

Based on the system testing results and observation, this research successfully developed a decentralized data storage network by implementing blockchain technology to prevent data alteration. The adoption of a transparent and decentralized system, with the use of blockchain technology as a distributed ledger, effectively addresses the inherent challenges posed by centralized data storage systems. This approach not only enhances the security, integrity, and availability of data but also presents a viable solution to the limitations associated with traditional systems. The results of the system testing unveil the prowess of the decentralized data storage system in navigating diverse scenarios, ranging from seamless addition and retrieval of data to the efficient acquisition of the latest block and retrieval of the complete blockchain. Notably, the system's adeptness in repairing an invalid node through the Gossip Protocol stands as a testament to its resilience and commitment to maintaining the integrity of data within the blockchain network.

Future research endeavors may chart a course toward scalability enhancements, optimization of consensus algorithms, and strategic considerations for network scaling and deployments. By delving deeper into these avenues, the research community can continue to advance the frontiers of decentralized data storage, ushering in an era of heightened data security, efficiency, and adaptability. In essence, this research not only contributes to the current understanding of decentralized systems but also paves the way for transformative innovations in the broader realm of data storage and management.

References

[1] Yoon-Jin Han, Geun-Hyung Kim, *"A Comparative Study on Decentralized Storage Platforms for Self-sovereign Data", doi: 10.21742/apjcri.2020.05.01, 2020.*

[2] Mohammed Ali Berawi, Mustika Sari, Fikroh Amalia Fahmi Addiani, Nunik Madyaningrum, *"Developing a Blockchain-based Data Storage System Model to Improve Government Agencies' Organizational Performance"*, Vol 12 No.5, doi: 10.14716/ijtech.v12i5.5237, 2021.

[3] Henderi, I Ketut Gunawan, Husni Teja Sukmana, Agung Yudo Ardianto, *"Blockhain Technology As A Me For Sharing Information That Generates User Access Rights and Incentives"*, Vol. 1 No. 1, doi: 10.34306/bfront.v1i01.2, 2021.

[4] Lu Meng, Bin Sun, *"Research on Decentralized Storage Based on a Blockchain"*, doi: 10.3390/su142013060, 2022.

[5] Wang Changjing, Jiang Huiwen, Zeng Jingshan, Yu Min, Huang Qing, Zuo Zhengkang, *"A Review of Blockchain Layered Architecture and Technology Application Research",* Vol.26 No.5, 415-428, doi: 10.19823/j.cnki.1007 1202.2021.0052, 2021.

[6] Kriti Bhalibar, Abhay Singh, Himani Sharma, Ashish Uphadyay, Himanshu Gupta, *"Centralize Storage System with Encryption vs Decentralize Storage System Using Blockchain"*, doi: 10.2139/ssrn.4119952, 2022.

[7] Mónica Martínez-Castañeda, Claudio Feijoo, *"Use of blockchain in the agri-food value chain: State of the art in Spain and some lessons from the perspective of public support"*, Vol.47. doi: 10.1016/j.telpol.2023.102574, 2023.

[8] E. Karaarslan, E.Konacaklı, *"Data Storage in the Decentralized World: Blockchain and Derivatives"*, doi: 10.26650/B/ET06.2020.011.03, 2020.

[9] Pranas Žukauskas, Jolita Vveinhardt and Regina Andriukaitienė, *"Exploratory Research"*, doi: 10.5772/intechopen.70631, 2018.

[10] Rama Aria Megantara, Farrikh Alzami, Ricardus Anggi Pramunendar, Dwi Puji Prabowo, *"Pengembangan dan Implementasi Docker Untuk Memaksimalkan Utilitas Server Universitas pada Masa Covid-19",* doi: 10.14710/transmisi.24.2.48-54, 2022.

[11] Fachru Dahri, Andi Marwan El Hanafi, Divi Handoko, Nur Wulan, *"Implementation of Microservices Architecture in Learning Management System E-Course Using Web Service Method"*, doi: 10.33395/sinkron.v7i1.11229, 2022.

[12] Eko Rudiawan Jamzuri, Hanjaya Mandala, Riska Analia, Susanto Susanto, *"Cloud-Based Architecture for YOLOv3 Object Detector using gRPC and Protobuf"*, doi: 10.15294/jte.v14i1.36537, 2022.

[13] J T Zhao, S Y Jing, L Z Jiang, *"Management of API Gateway Based on Micro-service Architecture"*, doi: 10.1088/1742-6596/1087/3/032032, 2018.

[14] Muhamad Rafli, *"Journal Web Server Load Balancing Performance Testing using Nginx Riverse Proxy Based on Centos 7 OS"*, Vol.9 No.3, doi: 10.35957/jatisi.v9i3.2185, 2022.

[15] Yan Zhu, Chunli Lv, Zichuan Zeng, Jingfu Wang, Bei Pei, *"Blockchain-based Decentralized Storage Scheme"*, doi: 10.1088/1742-6596/1237/4/042008, 2019.

[16] Santi Sulastri, Riana Defi Mahadji Putri, *"Implementasi Enkripsi Data Secure Hash Algorithm (SHA-256) dan Message Digest Algorithm (MD5) pada Proses Pengamanan Kata Sandi Sistem Penjadwalan Karyawan"*, doi: 10.15294/jte.v10i2.18628, 2018.

[17] Fariha Jahan, Mayel Mostafa, Shahrin Chowdhury, *"SHA-256 in Parallel Blockchain Technology: Storing Land Related Documents"*, doi: 10.5120/ijca2020920911, 2020.

[18] Untung Rahardja, Achmad Nizar Hidayanto, Ninda Lutfiani, Dyah Ayu Febiani, Qurotul Aini, *"Immutability of Distributed Hash Model on Blockchain Node Storage"*, doi: 10.15294/sji.v8i1.29444, 2021.

[19] A.S. Shaleva, V.V. Korkhov, *"Efficient Gosip-Based Protocol in The Neo Blockchain Network"*, 2021.

[20] Shadab Alam, *"T he Current State of Blockchain Consensus Mechanism: Issues and Future Works", doi: 10.14569/IJACSA.2023.0140810, 2023.*