Scientific Journal of Informatics
Vol. 8, No. 1, May 2021

p-ISSN 2407-7658 http://journal.unnes.ac.id/nju/index.php/sji e-ISSN 2460-0040

Immutability of Distributed Hash Model on Blockchain Node Storage

Untung Rahardja[1]*, Achmad Nizar Hidayanto[2], Ninda Lutfiani[3], Dyah Ayu Febiani[4], Qurotul Aini[5]
[1,3,5]Magister Program Informatics Department, Universitas Raharja, Indonesia
[2]Faculty of Computer Science, Universitas Indonesia, Indonesia
[4]Informatics Department, Universitas Raharja, Indonesia

Abstract
Purpose: The blockchain system uses hash functions. Hash is used in the blockchain to mark each block of data. Hash function algorithms map a string that is usually hexadecimal of any size to a sequence of fixed-size bits. This journal will discuss the distributed hash model for immutable blockchain node storage.
Methods: The methodology used in the preparation of this research is mind mapping and literature review, namely the collection of scientific journals, articles, and e-books.
Result: Storage of nodes on the blockchain using a distributed hash model. The distributed hash model only stores a portion of the block data at each node, and the block data is taken as a resource. A hash connects each block with the previous on the blockchain, so the entire blockchain transaction cannot be changed or deleted.
Novelty: For further research will expand the research topic regarding storage immutability on the blockchain so that it becomes more completed and detailed.

Keywords: Blockchain, Immutability, Distributed Hash, Node Storage, Block Data
Received March 2021 / **Revised** May 2021 / **Accepted** May 2021

This work is licensed under a Creative Commons Attribution 4.0 International License.

INTRODUCTION

The popularity of using digital currency, namely cryptocurrency for virtual transactions, is increasing. Bitcoin is the most widely used cryptocurrency [1]. Bitcoin uses a blockchain system for storage of transactions. Digital currency transactions can take place safely because of the blockchain system. Blockchain technology is registered in Bitcoin cryptocurrency transactions, the goal is to allow users to make transactions directly and be able to eliminate third party intermediaries [2].

Blockchain is an open distributed ledger which can record transactions between two parties efficiently and in a verifiable and permanent manner [3]. Blockchain usually uses a peer-to-peer system or from sender to receiver. In a peer-to-peer network, transactions or communications occur between one party and another without going through an intermediary. Each node is capable of storing and forwarding information to other nodes.

Blockchain technology continues to be used, so there are problems in storage. Bitcoin blockchain is the most commonly used blockchain system, as of February 1, 2021, the number of Bitcoin blocks was 668,618, and the blockchain size was 378.10 GB [4], and the number will continue to increase [5].

Because blockchain technology requires that each node stores all block data completely, which puts great pressure on the node storage system in the blockchain network and seriously limits the increase in computer access to the blockchain system [6].

Data security and confidentiality is very important in data exchange. The blockchain uses a hash model for node storage. Hash is a technique for converting an original message (plaintext) into random code to avoid being hacked. Hash functions convert any text or file into text strings of the same length. Any given input will still produce the same output length.

*Corresponding author.
Email addresses: untung@raharja.info (Rahardja), nizar@cs.ui.ac.id (Hidayanto), ninda@raharja.info (Lutfiani), dyah.ayu@raharja.info (Febiani), aini@raharja.info (Aini)
DOI: 10.15294/sji.v8i1.29444

Blockchain technology can save time and money in the transaction process and blockchain which has decentralized characteristics, namely a system that gives everything to all block chain users and the absence is prioritized, because they all have the same rights [7].

METHODS

The discussion of this paper is to analyze the immutability of blockchain node storage using a distributed hash model, so that the blockchain node storage is immutable or irreversible because it uses a hashing function. And the methodology used in the preparation of this research is mind mapping and other methods literature review, that is: Collection of scientific journals, articles and e-book collection. The data collection is a reference from open journals contained in online journals.

The research carried out is by using literature review, conducting studies on blockchain, and technology related to blockchain. An analysis of the hash function is carried out to determine the role and use of hashes in blockchain storage. How does the hash function work, what makes storage on the blockchain immutable. So that, it can find out how the hash function works in the immutable blockchain storage.

Analyze the use of hashing algorithms on blockchain data. Utilization of hash functions used in blockchain technology. A system of hash functions that can make the blockchain secure.

After analyzing the hash function on the blockchain, it can be explained about the results and discussion of the role and workings of the hash function on the blockchain node storage.

Mind Mapping

In preparing this study, the authors used the mind mapping method. Mind mapping is a method or technique used for learning material in a school or campus environment. According to Tony Buzan, "Mind mapping is a way to develop one's thinking activities from one direction to all directions, and can take various types of thoughts from different points of view".

Mind mapping can help plan a target, train communication, solve a problem quickly and accurately, make more focus, understand a concept or material as a whole, make it easier to determine priority tasks, can increase creativity, and understand the concept or problem that is complex [8]. Mind mapping blockchain can be seen in Figure 1.

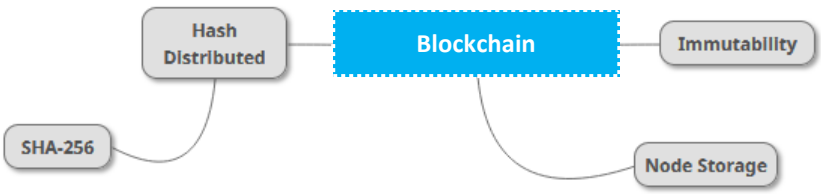

Figure 1. Mind Mapping Blockchain

This journal discusses an explanation of what blockchain is, the hash function used in this technology, blockchain node storage, and also an explanation of what makes the blockchain node immutable once it is created.

Blockchain technology has an advantage in terms of security [9], because blockchain is a ledger that records open transactions and uses a decentralized database that is spread around the world without going through third-party intermediaries. Initially blockchain was only used for encrypted digital currency transactions such as bitcoin, but research on digital currencies is being followed up, blockchain technology is

increasingly being developed, this technology deals with existing technologies such as network topologies, cryptography, and consensus algorithms, and is not just for making transactions [10].

Blockchain can be assumed as an archive of transactions collected on blocks with a time stamp. Each block is also identified by a hash value. However, each of these blocks references the hash value of an existing block [11]. Block Number Blockchain is presented in Figure 2.

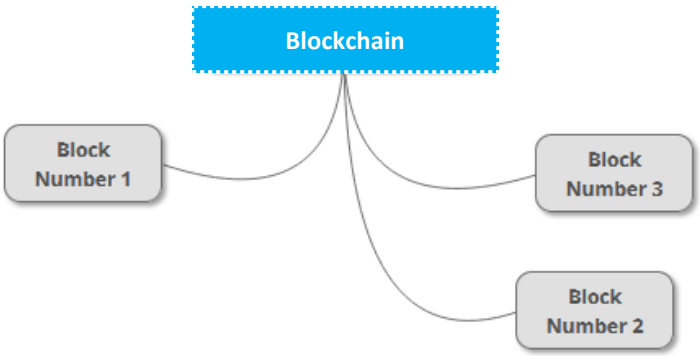

Figure 2. Block Number Blockchain

The blockchain system consists of two records, blocks, and transactions. Each block that makes up the network is made up of a cryptographic hash. Cryptographic systems are applied to secure the data sent so that the data can only be accessed by the sender and receiver. Cryptography is the science and art of storing messages securely when they are sent from one place to another [12].

The way blockchain works is when new data is stored in a block, and inside that block it uses a hash function that identifies the block and its contents in the form of a unique code. The blockchain is formed from previous blocks which carry historical information in each block. Blockchain records information or transactions that cannot be changed. Once the blockchain receives new information, the cryptographic hash takes that data and converts it into a unique code. This unique code can detect fraud. So that all transactions using blockchain become safer and more transparent. When making a transaction, there is public access without having to log in first, so the transaction can be seen by everyone.

Blockchain has character, that is, data that has been created cannot be changed (immutable) and can only be added (append only), it is a distributed ledger, all data is copied to network participants (nodes), and the data contained in the blockchain are connected to each other, so if someone wants to change a data in the node that cannot be done because they have to change the previous data.

Literature Review

Blockchain

In 2008 Satoshi Nakamoto introduced Blockchain via Bitcoin in the cryptocurrency field [1]. Blockchain is a system for recording transaction activities in many distributed databases. These transaction records can hardly be altered or hacked unilaterally because they are decentralized. Records of these transactions are contained in interconnected blocks. When a block is full, it will then create a new block connected to the previous block. If a block has been created, it means that the block cannot be changed anymore, this is called immutable.

Blockchain immutability or irreversibility states that transaction data in the education sector residing on the blockchain has never been tampered with, this historical data cannot be deleted or mutated Sunarya et al. (2020) [13].

Immutability is one of the characteristics of blockchain. Immutability is defined as immutable or also in the form of permanence. Immutable is the value of a data or variable that cannot be changed in program

execution. Blockchain is immutable which means it cannot be modified. This is caused by cryptographic hash functions. One of the cryptocurrencies, bitcoin uses SHA-256 (Secure Hash Algorithm 256 bits) created by the NSA. SHA-256 outputs a fixed length output of 265 bits. SHA-256 takes input data and generates a hash that is 256 bits or 64 character long.

In a blockchain, the output value, known as a hash, is used as a unique identifier for a block of data. The hash of each resulting block matches the hash of the previous block. What's more, the block hash is dependent on the data contained in the block, which means that any changes to that data will require a change to the block hash.

Therefore, each block created on the hash is based on the data contained in that block and the hash of the previous block. These hash markers play an important role in ensuring the security and immutability of the blockchain.

Server users cannot duplicate and falsify data because of the different transcripts in each block. The use of blockchain technology can manage the collected data well and the data is not easily faked [14].

One of the other concepts of blockchain technology is transparency, this concept allows the identity of the owner to be kept secret using cryptographic techniques so that the identity is encrypted, and the owner's identity is represented only by a hash [15].

Hash Function

Hash function or also known as one-way function (one-way-function), message digest, fingerprint, compression function, and message authentication code (MAC) [16]. A hash function is a mathematical function that accepts input of various sizes into output, usually a fixed length hexadecimal. Hash is usually written with a combination of numbers (0 to 9) and letters (a to f). Hash value or message digest is the output of the hash function [17].

The hash connects each block with the previous block on the Blockchain. That way, the entire Blockchain transaction cannot be changed or deleted. So, this makes the Blockchain safe from hacking.

In direct note, only the hash is stored to save space, rather than storing the assets or note on the blockchain. In blockchain the number of nodes allows for higher energy and storage costs, but the security is also higher [18].

The blockchain data structure is immutable, and can only be added. Every data from this Blockchain is connected to one another, if there is a change in one of the data blocks it will affect the next data.

Peer-to-Peer

Peer-to-peer (P2P) network is a network model consisting of two or more devices, where each station or computer contained in the network environment can share. This network makes it easy for users to transact directly without requiring a third party such as a bank [19]. Transactions are carried out peer-to-peer or from sender to recipient. All transaction records reside on computers on the network.

A decentralized network or peer to peer network on the blockchain will store various blocks of data which will later be connected to each other [1]. In a peer-to-peer system there is no central authority, so if one of the peers on the network is out there still has many peers to download from. Does not conform to the idealistic standards of the central system, so it is not susceptible to sensors.

Distributed Network

A distributed network is a network whose work system spreads software, computer programming, and data on many computers, and is interdependent with one another. The purpose of this network is to share resources [20].

The advantage of using this network is to solve network congestion and reduce errors when sharing resources as the number of computers is increased, computers or servers are distributed in many places. The costs incurred in using this network are also more efficient.

RESULT AND DISCUSSION

The use of Blockchain technology makes data security problems resolved because the data that enters the Blockchain through the encryption process becomes random code and will not be the same. Then, the data is distributed to all nodes or users in the Blockchain environment [21].

Data exchange between users in each transaction can be known by everyone, and the contents of the transaction cannot be seen other than the user concerned. Each node will match data from existing values by communicating with each other and connecting with other nodes. The process of the hash function is shown in Figure 3.

Figure 3. Process of the hash function

A hash function is a fixed-length mathematical function that takes variable-length input and converts it into a binary sequence. This function is used in various security applications and internet protocols. Some examples of its use are for message authentication, digital signatures, and password storage [22].

Hash functions on the blockchain are used as a technique for securing and validating data. Transaction data that will be added first are packaged into a data block before being converted using a hash function. The hash function on the blockchain uses the hash value of the previous block to perform the new block hash calculation. Therefore, each block is connected to each other as in a chain where changes in data in one block will affect the next blocks. Verification of the validity of a data can also be easily done by comparing the hash values between blocks. Hash chain is a collection of data blocks that are interconnected by a hash function.

This is how the hashing process works, we will try to enter certain inputs using the hashing process using SHA-256 (Secure Hashing Algorithm 256). Result example for hashing process with a different word is shown in Table 4 below.

INPUT	HASH
Hi	3639efcd08abb273b1619e82e78c29a7df02c1051b1820e99fc395dcaa3326b8
Welcome	0e2226b5235f0ff94a276eb4d07a3bfea74b7e3b8b85e9efca6c18430f041bf8

Table 4. Hashing process with different word

In SHA-256, regardless of the short length of an input, the SHA output will always remain 256-bits long. This is especially important when handling large amounts of data and transactions. So basically, instead of remembering the input data which could be huge, simply remembering the hash code and keeping track of the others. [23].

Small changes to the SHA-256 input can change the Hash result, the hash output change will be very large. The Result example is shown in Table 5.

INPUT	HASH
Hallo	753692ec36adb4c794c973945eb2a99c1649703ea6f76bf259abb4fb838e013e
hallo	d3751d33f9cd5049c4af2b462735457e4d3baf130bcbb87f389e349fbaeb20b9

Table 5. Hashing process with different capital letters

Although only the case changes in the first alphabet of the input, the effect on the hash output is very large. This is an important hash function because this property refers to one of the highest qualities of the blockchain, namely immutability. In SHA-256 small changes result in very different hash values. Regardless of the input size, the output will always be a fixed 256-bit size (64 characters). In addition, both outputs will always be constant, however much these words are executed by the algorithm.

CONCLUSION

Blockchain is a system for recording transaction activities in many distributed databases. Blockchain is a decentralized database, so there are no intermediaries in a transaction. Blockchain is immutable which means it cannot be changed. This is caused by cryptographic hash functions. Hash function is a function that is useful for shrinking/compressing a long input string into a shorter output string. Blockchain uses a hash function, so the entire Blockchain transaction cannot be changed or deleted. The hash connects each block with the previous block in the blockchain. This makes the blockchain safe from being hacked. Advice from the author for further research, it is better to expand the research topic so that it is not only about the issue of storage immutability on the blockchain, so that the discussion becomes more completed and detailed.

ACKNOWLEDGEMENT

The author would like to give a special Thanks to Universitas Raharja, especially the Alphabet Incubator, and also the supervisor, lecturer, and all parties involved in helping to completed this research.

REFERENCES

[1] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System, " *Manubot*, 2019.

[2] Q. Aini, U. Rahardja, N. P. L. Santoso, and A. Oktariyani, "Aplikasi Berbasis Blockchain dalam Dunia Pendidikan dengan Metode Systematics Review," *CESS (J. Comput. Eng. Syst. Sci.*, vol. 6, no. 1, pp. 58–66, 2021.

[3] M. Iansiti and K. R. Lakhani, *The Truth About Blockchain*, Reading, MA: Harvard Business School Publishing, 2017. [Ebook] Available: hbr.org.

[4] "Bitcoin, Litecoin, Namecoin, Dogecoin, Peercoin, Ethereum Stats," 2021. [Online] Available: https://bitinfocharts.com [Accessed February 01, 2021].

[5] F. P. Oganda, U. Rahardja, Q. Aini, M. Hardini, and A. S. Bist, "Blockchain: Visualization of The Bitcoin Formula," *PalArch's J. Archaeol. Egypt/Egyptology*, vol. 17, no. 6, pp. 308–321, 2020.

[6] B. Yu, X. Li, and H. Zhao, "Virtual Block Group: A Scalable Blockchain Model with Partial Node Storage and Distributed Hash Table," *Comput. J.*, vol. 63, no. 10, pp. 1524–1536, 2020.

[7] Z. Fauziah, H. Latifah, X. Omar, A. Khoirunisa, and S. Millah, "Application of Blockchain Technology in Smart Contracts: A Systematic Literature Review," *Aptisi Trans. Technopreneursh.*, vol. 2, no. 2, pp. 160–166, 2020.

[8] "Mind Mapping: Pengertian, Manfaat, dan Cara Membuatnya - Glints Blog." [Online] Available: https://glints.com/id/lowongan/mind-mapping-adalah/#.YDfB5ugzYwx [Accessed February 25, 2021].

[9] U. Rahardja, S. Kosasi, E. P. Harahap, and Q. Aini, "Authenticity of a Diploma Using the Blockchain Approach," *Int. J.*, vol. 9, no. 1.2, 2020.

[10] E. Zhang, "Antshares Whitepaper 1.0," 2016. [Online] Available: https://github.com/AntShares/AntShares/wiki/Whitepaper-1.0. [Accessed March 1, 2016].

[11] H. F. Putra, W. Wirawan, and O. Penangsang, "Penerapan Blockchain dan Kriptografi untuk Keamanan Data pada Jaringan Smart Grid," *J. Tek. ITS*, vol. 8, no. 1, pp. A11–A16, 2019.

[12] D. Ariyus, *Pengantar Ilmu Kriptografi: Teori Analisis & Implementasi*. Yogyakarta: Penerbit Andi, 2008.

[13] S. Kosasi, "Karakteristik Blockchain Teknologi dalam Pengembangan Edukasi," *ADI Bisnis Digit. Interdisiplin J.*, vol. 1, no. 1, pp. 87–94, 2020.

[14] F. Agustin, Q. Aini, A. Khoirunisa, and E. A. Nabila, "Utilization of Blockchain Technology for Management E-Certificate Open Journal System," *Aptisi Trans. Manag.*, vol. 4, no. 2, pp. 133–138, 2020.

[15] E. P. Harahap, Q. Aini, and R. K. Anam, "Pemanfaatan Teknologi Blockchain pada Platform Crowdfunding," *Technomedia J.*, vol. 4, no. 2, pp. 199–210, 2020.

[16] R. Damanik, "Pengkodean Pesan Teks Dengan Proses Penerapan Algoritma Kriptografi Secure

Hash Algorithm (SHA)," *J. Inform. Kaputama*, vol. 1, no. 1, pp. 48–57, 2017.
[17] R. Munir, *Kriptografi*, Bandung: Informatika, 2006.
[18] B. S. Riza, "Blockchain Dalam Pendidikan: Lapisan Logis di Bawahnya," *ADI Bisnis Digit. Interdisiplin J.*, vol. 1, no. 1, pp. 41–47, 2020.
[19] T. D. Danella, "Bitcoin Sebagai Alat Pembayaran Yang Legal Dalam Transaksi Online," *J. Mhs. Fak. Huk. Univ. Brawijaya*, 2015.
[20] P. Balda and S. M. Garg, "Security Enhancement in Distributed Networking," *Int. J. Comput. Sci. Mob. Comput.*, vol. 4, no. 4, p. 761, 2015.
[21] Q. Aini, N. Lutfiani, F. Hanafi, and U. Rahardja, "Application of Blockchain Technology for iLearning Student Assessment," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 14, no. 2, 2020.
[22] W. Stallings, *Cryptography and Network Security: Principles and Practices*, New Jersey: Prentice Hall, 2003.
[23] "What Is Hashing? [Step-by-Step Guide-Under Hood Of Blockchain]". [Online] Available: https://blockgeeks.com/guides/what-is-hashing (Accessed February 11, 2021).