Impacts on Fork Rate of Changes of Block Size and Block Generation Time for Blockchain Scalability

Taishi Nakai
Kyoto University
Kyoto, Japan

Kazuyuki Shudo
Kyoto University
Kyoto, Japan

***Abstract*—Improving the scalability of blockchain is one of the most important challenges faced by blockchain technology. To enhance scalability, methods such as increasing the block size or decreasing the average block generation time can be used. However, both approaches increase the fork rate and negatively impact security. Previous research has not determined whether increasing the block size or decreasing the block generation time has a more detrimental effect on the fork rate. In this paper, we mathematically demonstrate that changes in the block generation time have a greater impact on the fork rate than changes in block size, and that this is due to delays proportional to geographical distance. Furthermore, through simulation, we show that changes in the block generation time have a greater effect on the fork rate than changes in block size.**

***Index Terms*—blockchain, fork, scalability, security**

I. Introduction

Blockchain is a form of distributed ledger technology, that was first introduced by a paper published by Satoshi Nakamoto in 2008 [1]. At the core of this technology is a data structure that chains records of decentralized transactions, maintained by a multitude of participants (nodes) on the network without a central administrator or intermediary. Notably, once a block is added to the blockchain, its content possesses immutability, meaning it cannot be altered later. This characteristic implies that tampering with the data is extremely difficult, ensuring the reliability of the blockchain.

In recent years, blockchain technology has been adopted in various industries and scenarios, from finance [2] to supply chain management [3], healthcare [4], real estate [5], and energy management [6]. The decentralized nature and tamper-resistance of this technology contribute to increased transparency and trustworthiness, leading many companies and organizations to consider its implementation. However, as blockchain adoption grows, scalability has emerged as a significant challenge.

The scalability issue of blockchain primarily relates to transaction processing capacity and speed. Many current blockchains have a limit to the number of transactions they can process within a given time, leading to potential delays or increased fees when there's a surge in transactions. This scalability issue poses a serious barrier, especially in sectors like financial institutions handling large-scale transactions or supply chains requiring real-time data processing.

To address this scalability issue, many researchers are working on various methods and approaches to enhance blockchain scalability. For instance, Zhou et al. [7] conducted a comprehensive survey of existing solutions to the blockchain scalability problem, categorizing them by levels. Kim et al. [8] analyzed various attempts to solve the blockchain scalability issue, categorizing them into On-chain, Off-chain, Side-chain, Child-chain, and Inter-chain, providing a detailed examination. These studies represent crucial efforts in seeking solutions to the blockchain scalability problem, laying the groundwork for the further evolution and proliferation of blockchain technology.

The simplest method to improve scalability is to increase the block size B or reduce the average block generation time T. This is because $\frac{B}{T}$ represents the data processing capacity of the blockchain per unit time, and increasing B or reducing T increase $\frac{B}{T}$. Indeed, Ethereum [9] improved its scalability by increasing the block size B and reducing the block generation time T. Furthermore, Bitcoin Cash [10], which hard-forked from Bitcoin, improved the scalability by increasing the block size B. Additionally, Bitcoin Candy [11], which hard-forked from Bitcoin Cash, enhanced the scalability by reducing the average block generation time T.

However, increasing the block size B or reducing the block generation time T can lead to forks, compromising security [12]. A fork occurs when two or more blocks are generated at the same height. A larger block size means longer propagation times, increasing the likelihood of simultaneous block generation on the blockchain network and, consequently, a higher chance of forks. Similarly, a shorter block generation time means more frequent block generation, again increasing the likelihood of forks. A high fork rate increases the risk of attacks like Double Spending Attack or Selfish Mining [13]. Ethereum addressed the deteriorating security from increasing B or reducing T by partially adopting the GHOST (Greedy Heaviest Observed Subtree) protocol [12].

However, it is unclear whether increasing the block size B or reducing the block generation time T has a more detrimental effect on the fork rate. Without a clear understanding of the factors negatively impacting the fork rate, it is challenging to determine the best optimization or improvement strategy for blockchain.

The contributions of this paper are as follows:

1) In blockchains that adopt Proof of Work, we mathematically demonstrate that changing the block generation time T has a larger effect on the fork rate than changing

Proc. IEEE 42nd Int'l Conf. on Consumer Electronics
(IEEE ICCE 2024), January 2024

the block size B, and show that the cause lies in the delay proportional to geographical distance. This also means reducing the block generation time T has a more detrimental effect on the fork rate than increasing the block size B.

2) Through simulations, we show that changing the block generation time T has a larger effect on the fork rate than changing the block size B. This also means that for scalability improvements, reducing the block generation time T has a more detrimental effect on the fork rate than increasing the block size B.

II. Mathematical Description

According to Nakai et al. [14] (accepted), the formula for the blockchain trilemma for Proof of work is as follows:

$$\frac{B_h + B_{tx} \cdot n_{tx}}{T} \cdot \frac{1}{F} \cdot \boldsymbol{H}^\top \boldsymbol{P} \boldsymbol{H} = 1 \tag{1}$$

where B_h is a block header size, B_{tx} is a transaction size, n_{tx} is the number of transactions in a block, T is the block generation time, and F is the fork rate. The vector $\boldsymbol{H}$ and matrix $\boldsymbol{P}$ are specifically represented as:

$$\boldsymbol{H} = \begin{pmatrix} H_1 \\ H_2 \\ \vdots \\ H_n \end{pmatrix} \tag{2}$$

$$\boldsymbol{P} = \begin{pmatrix} 0 & t_{12} & \dots & t_{1n} \\ t_{21} & 0 & \dots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ t_{n1} & t_{n2} & \dots & 0 \end{pmatrix} \tag{3}$$

$\boldsymbol{H}^\top$ represents the transposed vector of $\boldsymbol{H}$. n denotes the number of nodes participating in the network, and H_i is the proportion of the hash rate that node i possesses relative to the entire network's hash rate. t_{ij} is the time per byte it takes for a block generated by node i to propagate to node j. In matrix $\boldsymbol{P}$, the diagonal elements indicate the time it takes for block propagation between the same nodes, and since a block does not propagate between the same nodes, they are 0. Equation (1) indicates that $\frac{B_h + B_{tx} \cdot n_{tx}}{T}$ contains the scalability term, $\frac{1}{F}$ contains the security term, and $\boldsymbol{H}^\top \boldsymbol{P} \boldsymbol{H}$ contains the decentralization term. The product of these three terms equals 1, illustrating the trilemma.

Here, $\boldsymbol{H}^\top \boldsymbol{P} \boldsymbol{H}$ is a quadratic form and can be represented without using vectors and matrices as:

$$\boldsymbol{H}^\top \boldsymbol{P} \boldsymbol{H} = \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} t_{ij} H_i H_j \tag{4}$$

Substituting equation (4) into equation (1) gives:

$$\frac{B_h + B_{tx} \cdot n_{tx}}{T} \cdot \frac{1}{F} \cdot \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} t_{ij} H_i H_j = 1 \tag{5}$$

$B_h + B_{tx} \cdot n_{tx}$ represents the block size and is equal to B. Thus:

$$\frac{B}{T} \cdot \frac{1}{F} \cdot \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} t_{ij} H_i H_j = 1 \tag{6}$$

Let T_{ij} be the time it takes for a block generated by node i to propagate to node j. Using the data transfer capacity C, which indicates the amount of data that can be transferred per second, and the delay time D proportional to the geographical distance, T_{ij} has the following proportional relationship:

$$T_{ij} \propto \frac{B}{C} + D \tag{7}$$

The time t_{ij} per byte it takes for a block generated by node i to propagate to node j is represented by the following equation:

$$t_{ij} \propto \frac{1}{C} + \frac{D}{B} \tag{8}$$

When x is a value greater than 1, in equation (6), if T is multiplied by $\frac{1}{x}$, $\frac{B}{T}$ increases by x times, and F also increases by x times. On the other hand, if B is multiplied by x, t_{ij}, which is proportional to $\frac{1}{C} + \frac{D}{B}$, decreases. Therefore, in this case, while $\frac{B}{T}$ increases by x times, F increases by less than x times. Conversely, when x is a value less than 1, in equation (6), if T is multiplied by $\frac{1}{x}$, $\frac{B}{T}$ decreases by x times, and F also decreases by x times. However, if B is multiplied by x, t_{ij}, which is proportional to $\frac{1}{C} + \frac{D}{B}$, increases. Therefore, in this case, while $\frac{B}{T}$ decreases by x times, F decreases by less than x times. In other words, the increase or decrease in T has a greater impact on F than the increase or decrease in B.

From the above, it can be understood that reducing the block generation time T has a more detrimental effect on the fork rate than increasing the block size B, and the cause lies in the delay time D proportional to the geographical distance.

III. Simulation

We verify through simulation that reducing the block generation time T has a more detrimental effect on the fork rate than increasing the block size B.

For the simulation, we use Simblock [15]. The basic settings for the simulation are as follows:

- Number of nodes: 1,000
- Inbound connections: 125
- Outbound connections: 8
- Reconnect every 10 blocks
- Up to a maximum block height of 100,000

We set the block size B to 535,000 bytes and T to 600,000 milliseconds, and at this time, $\frac{B}{T}$ is 1 as the standard. By fixing either B or T and varying the other, we change the value of $\frac{B}{T}$ and measure the theoretical fork rate [16]. Specifically, we measure $\frac{B}{T}$ values of 0.0625, 0.125, 0.25, 0.5, 1, 2, 4, 8, and 16.

The simulation results are shown in Table I and Figure 1. In the range where $\frac{B}{T}$ is greater than 1, it can be seen that the line where T is varied has a larger increase in the fork rate

TABLE I

COMPARISON BETWEEN "FORK RATE WHEN CHANGING T AND FIXING B" AND "ONE WHEN CHANGING B AND FIXING T".

$\frac{B}{T}$	0.0625	0.125	0.25	0.5	1	2	4	8	16
F when Changing B and Fixing T	0.00111	0.00132	0.00178	0.00274	0.00480	0.00869	0.01639	0.03225	0.06250
F when Changing T and Fixing B	0.00025	0.00063	0.00114	0.00239	0.00480	0.00943	0.01886	0.03846	0.07692

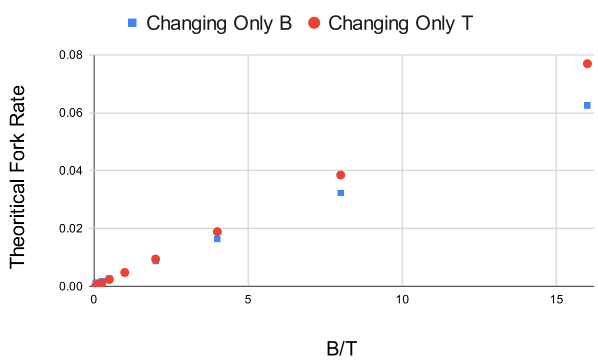

Fig. 1. The change in T has a stronger impact on the theoretical fork rate than the change in B.

than the line where B is varied. Also, in the range where $\frac{B}{T}$ is less than 1, it can be seen that the line where T is varied has a larger decrease in the fork rate than the line where B is varied. This shows that the block generation time T has a greater impact on the fork rate than the block size B. It also means that reducing the block generation time T has a more detrimental effect on the fork rate than increasing the block size B.

IV. CONCLUSION

We mathematically demonstrated that changing the block generation time T has a greater effect on the fork rate than changing the block size B, based on the blockchain trilemma equation and the fact that the time it takes for a block generated by node i to propagate to node j, T_{ij}, is proportional to $\frac{B}{C}+D$, and that the cause of this is the delay time proportional to the geographical distance. This means that reducing the block generation time T has a more detrimental effect on the fork rate than increasing the block size B. Furthermore, through simulation experiments, we demonstrated that changing the block generation time T has a greater effect on the fork rate than changing the block size B. This also means that reducing the block generation time T has a more detrimental effect on the fork rate than increasing the block size B.

In this study, we focused on blockchains that adopt Proof of Work. We plan to verify whether reducing the block generation time T has a more detrimental effect on the fork rate than increasing the block size B in blockchains that adopt Proof of Stake [17].

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number JP21H04872.

REFERENCES

[1] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. 2008.

[2] Victor Chang, Patricia Baudier, Hui Zhang, Qianwen Xu, Jingqi Zhang, and Mitra Arami. How blockchain can impact financial services – the overview, challenges and recommendations from expert interviewees. *Technological Forecasting and Social Change*, 158:120166, 2020.

[3] Rosie Cole, Mark Stevenson, and James Aitken. Blockchain technology: Implications for operations and supply chain management. *Supply Chain Management : an International Journal*, 24(4):pp 469 – 483, 2019.

[4] Asaph Azaria, Ariel Ekblaw, Thiago Vieira, and Andrew Lippman. Medrec: Using blockchain for medical data access and permission management. In *2016 2nd International Conference on Open and Big Data (OBD)*, pages 25–30, 2016.

[5] Avi Spielman. Blockchain : digitally rebuilding the real estate industry. 2016.

[6] Arzoo Miglani, Neeraj Kumar, Vinay Chamola, and Sherali Zeadally. Blockchain for internet of energy management: Review, solutions, and challenges. *Computer Communications*, 151:395–418, 2020.

[7] Qiheng Zhou, Huawei Huang, Zibin Zheng, and Jing Bian. Solutions to scalability of blockchain: A survey. *IEEE Access*, 8:16440–16455, 2020.

[8] Soohyeong Kim, Yongseok Kwon, and Sunghyun Cho. A survey of scalability solutions on blockchain. In *2018 International Conference on Information and Communication Technology Convergence (ICTC)*, pages 1204–1207, 2018.

[9] Vitalik Buterin et al. A next-generation smart contract and decentralized application platform. *white paper*, 3(37):2–1, 2014.

[10] Marco Alberto Javarone and Craig Steven Wright. From bitcoin to bitcoin cash: A network analysis. In *Proceedings of the 1st Workshop on Cryptocurrencies and Blockchains for Distributed Systems*, CryBlock'18, pages 77–81, New York, NY, USA, 2018. Association for Computing Machinery.

[11] Bitcoin candy whitepaper, 2018. https://cdy.one/whitepaper.pdf, (Accessed on 10/23/2023).

[12] Yonatan Sompolinsky and Aviv Zohar. Secure high-rate transaction processing in bitcoin. In Rainer Böhme and Tatsuaki Okamoto, editors, *Financial Cryptography and Data Security*, pages 507–527, Berlin, Heidelberg, 2015. Springer Berlin Heidelberg.

[13] Arthur Gervais, Ghassan O. Karame, Karl Wüst, Vasileios Glykantzis, Hubert Ritzdorf, and Srdjan Capkun. On the security and performance of proof of work blockchains. In *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*, pages 3–16, 2016.

[14] Taishi Nakai, Akira Sakurai, Shiori Hironaka, and Kazuyuki Shudo. The blockchain trilemma described by a formula. In *IEEE Blockchain*, 2023. (accepted).

[15] Yusuke Aoki, Kai Otsuki, Takeshi Kaneko, Ryohei Banno, and Kazuyuki Shudo. Simblock: A blockchain network simulator. In *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 325–329, 2019.

[16] Akira Sakurai and Kazuyuki Shudo. Impact of the hash rate on the theoretical fork rate of blockchain. In *IEEE ICCE 2023*, pages 1–4. IEEE, 2023.

[17] David Bernardo Kiayias Aggelos, Russell Alexander and Oliynykov Roman. Ouroboros: A provably secure proof-of-stake blockchain protocol. In Katz Jonathan and Shacham Hovav, editors, *Advances in Cryptology – CRYPTO 2017*, pages 357–388, Cham, 2017. Springer International Publishing.