EasyChair Preprint

№ 9086

How Efficient is Blockchain-Based Data Analyser for Supply Chain?

Zhile Yu

October 24, 2022

How Efficient Is Blockchain-Based Data Analyser For Supply Chain?

Zhile Yu
Faculty of Information Technology
Griffifth University
Gold Coast, Australia
zhile.yu@griffithuni.edu.au

***Abstract*— A highly developed civilization needs an efficient and resolute approach to executing and solving problems by communicating data within the supply chain. Today, a blockchain that applies concepts such as peer-to-peer, nonce, and digital signature, is renowned to be the most preferred logistical method. The development of blockchain has gained a lot of attention from small to large-scale companies. Walmart and Carrefour have seized the opportunity very well and greatly increased their revenue. By using a blockchain-based data analyzer, businesses intended to make use of information to bring efficiency to upcoming transactions. Evidence shows that technology can address the issues of low efficiency, high cost, and repetitive work. However, there are challenges to implementation due to the exclusion of off-chain data in the blockchain as well as the limited scalability when there is a flooding of transactions. Nevertheless, solutions have been initiated by many researchers to mitigate these challenges, which include but are not limited to authentication of the transaction at execution time and implementing sharding. A successful supply chain will bring the organization a great way to manage its products.**

Keywords—blockchain, supply chain, data analyzer, efficient, communication, business

I. INTRODUCTION

The first emergence of blockchain in distributed ledger technology in 2008 has significantly impacted the lives of many individuals. It has shifted from the traditional way of communication and making offline transactions to making use of digital networking and the growing system of blockchain. A blockchain is a chain consisting of one block after another according to the time sequence of the transaction, in which each block stores a certain amount of information [1]. It upholds the virtues of integrity, confidentiality, and authentication, which endorses strong transparency, honesty, and security for the stakeholders [2]. On these grounds, information can easily be traced, financial transactions become more reliable and authentic, and potentially aids in solving many supply chain problems related to the cost of data flow and interaction and adaptability [1]. It is believed that blockchain can be an evolving method to record and track supply chain transaction that connects from one stakeholder to another. Also, it acts as the foundation for a data analyser tool that empowers predictive analytics, analyses real-time data, and manages data sharing. This report will discuss to what extent blockchain-based data analyser for supply chain effective and efficient, by examining applied blockchain concepts and various aspects such as interoperability, scalability, security, and challenges, alongside the solutions to minimise risks of the imposed negative effects.

II. KEY CONCEPTS OF BLOCKCHAIN

A. Peer-to-Peer

Peer-to-peer, abbreviated as P2P, is defined as one or more computers that connect, and share resources without going through a separate server computer [3]. An advanced P2P network allows data sharing and exchange to become much easier, accessible, and faster even over large distances [4]. It is no longer required to rent a building or invest in a separate server computer, which reduces costs incurred. Moreover, P2P helps to build trust between the stakeholders because transactions can take place without a third party being involved. Also, it has a better security system compared to the traditional, centralized client-server system. Due to decentralization, even if one node has a problem, the other nodes can still be maintained, and the whole system will not be tampered with [5]. Furthermore, it has a relatively high degree of resistance to malicious activity, especially since it is immune to Denial-of-Service (DoS) attacks [3]. However, a trade-off to its increased level of security, adding one transactional data require an update on every single node in the distributed ledger of the blockchain, which incurs an extensive amount of computational time complexity [4].

B. Nonce

The nonce is an abbreviation for “number only used once”. It is the number that the blockchain miners are trying to solve [6]. It is a random number that is added to a hashed block in a blockchain, with the purpose of when it is rehashed, it should meet the difficulty level restrictions [7]. Back to the idea of blockchain being applied in the supply chain, the Nonce enhances security because it efficiently and successfully prevents the hasher to attack the system. In many applications, hashing is done three times to a document, to give extra protection so that hacker finds it very difficult to decrypt the original document [6]. Moreover, reference [3] shows some arguments by Li, Barenji, and Huang that nonce specifies the order of transaction and originality of data, in which stakeholders can easily know which block is deriving from which transaction. It is also useful to track and manage the history of transactions. Furthermore, nonce also helps prevent double spending attacks, which is an attack that is caused by issues in 50% of nodes and makes a new block to deceive and manipulate the system [6]. A nonce can easily solve the issue of a double sending attack because Nonce has a one-time mechanism, which the nonce can only be used once.

C. Digital signature

A digital signature is a way of protecting documents that utilize a mathematical algorithm to encrypt documents [8]. It uses a hash function for asymmetric encryption to verify the authenticity of the document and the integrity of both the sender and the receivers [8]. For example, the sender sends the document to the receiver, the document will be encrypted once, and then the receiver hashes the document to verify its authenticity. They are a fundamental building block in blockchain-based data analyzer, as it is used mainly to authenticate transactions that includes information such as transaction time, types, and amount to authenticate it [8].

III. Discussion On Various Aspects

A. Interoperability

In the blockchain, interoperability is regarded as the ability to interact, share, and exchange data freely across different blockchain networks [9]. Interoperability allows to connect systems and helps to provide a broad view for analysis, which provides a completed report to stakeholders [3]. Due to this aspect, blockchain directly reduces the cost of information interaction. When a supplier or buyer needs to make a transaction, they no longer need to seek a third party to maintain the transaction since a blockchain-based data analyzer can support it. Walmart, for example, is linked to farms through blockchain and supply chain applications [3]. It is the data and information stored in each block, and it continually comes from the first stage farmer to the last stage customer purchasing the products which the farmer produces, as can be seen in figure 1. This greatly shortens the complex and costly processes of going through middlemen at Walmart and farms, so it has been a lot of success [9]. Blockchain-based data analyser enables the information that interacts in each block to increase the reliability of upcoming transactions because it is conducted by machine learning.

Fig. 1. Data flow within the supply chain.

B. Scalability

The scalability of blockchain refers to the ability and capacity of the network to support a larger number of transactions and nodes and how much chain and block it can extend [10]. This aspect is critical for the future growth of blockchain. Reference [11] was the study conducted by Khan, Jung, and Hashmani, that evident the size of the distributed ledger and the number of transactions considerably grow at a constant rate. This signifies that a blockchain-based data analyzer is efficient and sustainable because it can contain and adjust when there is an anticipated increasing volume of transactions. Also, blockchain scalability offers data security, autonomy, immutability, and transparency benefits with the aid of an increasing number of nodes in the network [11]. However, there is a trade-off between scalability and performance, such that it is difficult to manage and control all the nodes, blocks, and chains in a short period [10]. Also, with the large scalability and capacity of data, it incurs storage space to separate and distribute the information into blocks and nodes in a chain [11].

C. Security

Without centralized server computers, people use peer-to-peer to share resources alternatively. The system is relatively more secure because the separate and distributed data into different nodes become more difficult to attack. Whilst there are still chances to be attacked by hackers, the attacks become more visible and noticeable, which allows them quicker to be resolved [12]. Also, referring to the concept of decentralization, every node has its significant role, however, if one node is down, the others can pick up the slack, so that the overall network still operates continuedly and is not affected completely [12]. In addition, the distributed network has over ten thousand nodes all over the world that keep track of all transactions happening on the system [13]. From the perspective of social development, blockchain technology is still in the early stage of its development. There is a high urgency for laws to be enforced, and governments and social groups to make interventions and contributions to the recognition of data safety and security and pay attention to social acceptance, and security issues of legal terms [13].

IV. Significance Of Blockchain

Blockchain-based data analyzer is found to significantly improve the cost-efficient delivery of products, enhancing products' traceability, improving coordination between partners, and aiding access to financing [14]. It can make the whole system work smoother and the information from the data analyzer can make the next service for users more efficient. The blockchain-based data analyzers can survey the experience form user of the decentralized mechanism.

For example, Walmart, the leader in supply chain management, has successfully used decentralized distributed-ledger technology to create an automated process for handling invoices and payments to seventy external freight carriers since 2017 [15]. The decentralized mechanism offers safe storage and a reliable algorithm, which allows a transparent and dependable environment for different stakeholders. Another example, pork products in the farms are set up with a QR code on each box, in which the details are uploaded into the blockchain-based data system and each node can confirm the details of the operation [15]. This allows each food steward to easily receive and access information on the pork products at a fast-paced, which is time-efficient and cost-efficient. Reference [16] was the paper written by Azzi, Chamoun, and Sokhn stating by using this technology, stakeholders can quickly and easy to solve issues when they find faulty products. It also reduces time consumption and prevents someone unauthorized from tampering with information.

Similarly, Carrefour also uses a blockchain-based data analyzer to record information on products such as chicken and vegetables since 2020 [15]. Due to the COVID-19 outbreak with many cities being constantly locked down, the

logistic links between cities and countries continue to decline [15]. Carrefour has launched a plan to improve its supply chain management by using blockchain. Almost all of Carrefour's stakeholders support this plan because traceability and food safety of the products play the greatest significance [15]. Also, the efficiency of transportation, logistics, and supply chains has been greatly reduced during the epidemic, which are important forces that can quickly recover from the epidemic [15].

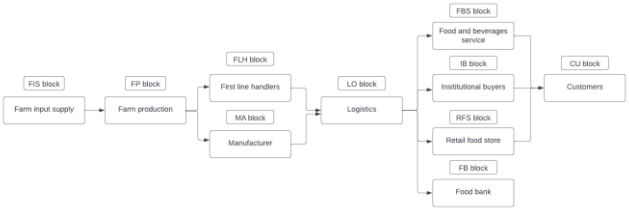

Fig. 2. Process of implementation.

The figure above portrays an example of the flow of data when implementing a blockchain-based data analyser into the supply chain in the food industry. The process begins at the Farm as the initial input, which is connected to the second conjunction farm production. The farm production is linked to first-line handlers and manufacturers. From the first line header and manufacturer, logistics connects food and beverage service, institutional buyers, retail food stores and food banks, in this stage, shows the lower supply chain stream and it more in customer service part. From the retail store, it all comes to customers and customers refer to every stage of the supply chain. This process of implementation clearly shows that the stakeholders and organization can successfully share and exchange data without the need for a middleman. This methodology brings significant benefits to stakeholders such as traceability, information transparency, efficient food distribution, and improved food safety and quality.

V. Challenges

A. *Off-Chain Data Are Not Maintained*

This technology is notable for its excellent performance in many aspects of security and information transparency. However, it is limited to only maintaining security features for on-chain data. Off-chain data are any non-transactional data external to a blockchain, usually unstructured and in any form of image or text [17]. It includes laws and regulations that appear along with the transaction, the authenticity of checks, and the authenticity of off-chain orders or accounts receivable claims that are not regulated [9]. It is a challenge to maintain off-chain data because they commonly require greater storage to store the large data size for each off-chain data item [5]. What is stored in the blockchain is the hash or digital signature generated for each off-chain data, not the data item due to the limited capacity of storage. The actual off-chain data item may be stored in cloud storage, which is expected since it will exceed the needs of the blockchain storage [17]. Also, with newer development and implementation in the blockchain-based data analyser that off-chain data will be maintained, another challenge arises because blockchain data are unchangeable. This is a problem when privacy laws require companies to delete data permanently once they have served its purpose and is no longer needed for documentation [9]. Other issues that may arise are related to the accessibility and transaction speed since each off-chain data will require additional data connectors, encryption, and security application [17].

B. *Flooding of Transactions and Limited Scalability*

Another limitation of the blockchain-based data analyzer is dealing with limited scalability. It is potentially a challenge because it requires sufficient space to support the large capacity of the blockchain system. The growing number of nodes has huge management risks, and the development of blockchain often cannot avoid the continuous expansion of the blockchain system itself [10]. The increasing number of nodes is directly proportional to the number of risk factors it sustains [18]. Also, it is often the case that larger blockchain systems have greater management challenges that cannot be avoided [18]. The transaction output and transaction confirmation have latency problems such that it is difficult to handle and manage large data that are separated and distributed into blocks in a short period [10]. P2P transactions, the broadcasting or flooding of all transactions and block creation, make the availability of bandwidth of the network a bottleneck.

VI. Solution To Mitigate Challenges

A. *Authenticate Off-Chain Data At Execution Time*

Blockchain indeed addresses most of the security problems for data, especially with aid of smart contracts and digital signatures using algorithms. However, there is no easy way to verify the authenticity of the transactional data simultaneously. One practical method proposed is to notarize the data at its execution time. This means that for every transaction, it is required to confirm the previous state of the corresponding transaction first [1]. For authenticating one data item, many algorithms can be applied to verify the authenticity. Among the many algorithms, a hash function is cheap, practical, and has a high safety factor, which is also economical and practical and widely used due to its mature performance [3]. After the calculation of Sha256, a 256-bit binary representation will be generated, that is the expression of 256/4 = 64-bit Hexadecimal [8].

Another method that can mitigate issues related to accessibility and transaction speed is to build a shared network of server and storage resources that are intended to give every block and node the necessary environment and security [13]. The identity of each retrieved data item must be established using previously saved hash values, demonstrating that it is the same item as the one that was initially stored. To ensure that the decentralised system works, in which the loss of one node does not result in major data loss while the node is down, each data item should be kept in many data storages [19]. Additionally, a system to synchronize off-chain references and rebalance the off-chain

data are needed once a node rejoins the chain following recovery [19].

B. Sharding For Improved Scalability

One prominent on-chain scalability technique is sharding. It focuses on allowing the blockchain network to split into smaller, more manageable, compact networks [20]. It is more effective for data processing when each shard takes a data item, as this truncates to each other's [20]. The concept of traditional database sharding, in which the database is divided into portions and spread across several servers, serves as a model for blockchain sharding. In a network that is sharded, nodes are split up into different shards, and network transactions are also split up into different shards [20]. As a result, only a small part of incoming transactions is handled by each node, while the remainder is handled simultaneously by other nodes in the network. More transactions could be completed and confirmed simultaneously by fragmenting the network. Many blocks can be checked simultaneously when only one block can be confirmed at once, completing the throughput expansion. Sharding will significantly accelerate transaction processing [20].

VII. Conclusion

In conclusion, blockchain technology has several applications, including transactions, data exchange, and trust establishment. Particularly, the supply chain industry will profit greatly from native components that the blockchain-based data analyzer will deliver. Peer-to-peer, nonce, and digital signature concepts are the foundation of blockchain, which bring effects to various aspects including interoperability, scalability, and security. It was discovered through analysis of real-world applications that there are various challenges when using blockchain-based data analyzers. These include the exclusion of off-chain data in the blockchain and limited capability due to flooding of transactions. These two issues make the applications of blockchain in the supply chain less effective, making data difficult to manage and forecast, and question their validity. These two problems can be relieved by sharding for increased scalability and authenticating off-chain data at execution time. While these solutions may not be able to eliminate the problems, they can nonetheless significantly cut costs and workload. Therefore, the effectiveness of blockchain-based data analysers for the supply chain can always be attributed to a strong beginning to allow new models alongside with critical evaluation of risks, difficulties, and public acceptance to demonstrate its distinctive problem-solving capabilities.

References

[1] U. Agarwal et al., "Blockchain Technology for Secure Supply Chain Management: A Comprehensive Review", *IEEE Access*, vol. 10, pp. 85493-85517, 2022. Available: 10.1109/access.2022.3194319.

[2] Y. Zou, T. Meng, P. Zhang, W. Zhang and H. Li, "Focus on Blockchain: A Comprehensive Survey on Academic and Application", *IEEE Access*, vol. 8, pp. 187182-187201, 2020. Available: 10.1109/access.2020.3030491.

[3] Z. Li, A. Barenji and G. Huang, "Toward a blockchain cloud manufacturing system as a peer to peer distributed network platform", *Robotics and Computer-Integrated Manufacturing*, vol. 54, pp. 133-144, 2018. Available: 10.1016/j.rcim.2018.05.011.

[4] S. King and S. Nadal, 2012. [Online]. Available: https://www.semanticscholar.org/paper/PPCoin%3A-Peer-to-Peer-Crypto-Currency-with-King-Nadal/0db38d32069f3341d34c35085dc009a85ba13c13. [Accessed: 30- Sep- 2022].

[5] D. Vangulick, B. Cornélusse and D. Ernst, "Blockchain for Peer-to-Peer Energy Exchanges: Design and Recommendations", *Ieeexplore.ieee.org*, 2022. [Online]. Available: https://ieeexplore.ieee.org/document/8443042/. [Accessed: 30- Sep- 2022].

[6] D. MacKenzie, *Pick a nonce and try a hash*, 41st ed. London: London Review of Books, 2019.

[7] L. Ismail, H. Hameed, M. Alshamsi, M. Alhammadi and N. Aldhanhani, "Towards a Blockchain Deployment at UAE University: Performance Evaluation and Blockchain Taxonomy", *ResearchGate*, 2019. [Online]. Available: https://www.researchgate.net/publication/333044281_Towards_a_Blockchain_Deployment_at_UAE_University_Performance_Evaluation_and_Blockchain_Taxonomy. [Accessed: 30- Sep- 2022].

[8] A. Khalique, K. Singh and S. Sood, "Implementation of Elliptic Curve Digital Signature Algorithm", *International Journal of Computer Applications*, vol. 2, no. 2, pp. 21-27, 2010. Available: 10.5120/631-876.

[9] S. Schulte, M. Sigwart, P. Frauenthaler and M. Borkowski, "Towards Blockchain Interoperability", 2022. .

[10] G. Karame, "On the Security and Scalability of Bitcoin's Blockchain | Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security", *ACM Conferences*, 2022. [Online]. Available: https://dl.acm.org/doi/10.1145/2976749.2976756. [Accessed: 30- Sep- 2022].

[11] D. Khan, L. Jung and M. Hashmani, "Systematic Literature Review of Challenges in Blockchain Scalability", *Applied Sciences*, vol. 11, no. 20, p. 9372, 2021. Available: 10.3390/app11209372.

[12] R. Zhang, R. Xue and L. Liu, "Security and Privacy on Blockchain | ACM Computing Surveys", *ACM Computing Surveys*, 2022. [Online]. Available: https://dl.acm.org/doi/10.1145/3316481. [Accessed: 30- Sep- 2022].

[13] C. Ye, G. Li, H. Cai, Y. Gu and A. Fukuda, "Analysis of Security in Blockchain: Case Study in 51%-Attack Detecting", *Ieeexplore.ieee.org*, 2022. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8563187. [Accessed: 30- Sep- 2022].

[14] J. Xu et al., "Healthchain: A Blockchain-Based Privacy Preserving Scheme for Large-Scale Health Data", *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8770-8781, 2019. Available: 10.1109/jiot.2019.2923525.

[15] B. Tan, J. Yan, S. Chen and X. Liu, "The Impact of Blockchain on Food Supply Chain: The Case of Walmart", 2022. .

[16] R. Azzi, R. Chamoun and M. Sokhn, "The power of a blockchain-based supply chain", *Computers & Industrial Engineering*, vol. 135, pp. 582-592, 2019. Available: 10.1016/j.cie.2019.06.042.

[17] S. Chang and Y. Chen, "When Blockchain Meets Supply Chain: A Systematic Literature Review on Current Development and Potential Applications", *IEEE Access*, vol. 8, pp. 62478-62494, 2020. Available: 10.1109/access.2020.2983601.

[18] Q. Zhou, H. Huang, Z. Zheng and J. Bian, "Solutions to Scalability of Blockchain: A Survey", *IEEE Access*, vol. 8, pp. 16440-16455, 2020. Available: 10.1109/access.2020.2967218.

[19] W. Reijers, I. Wuisman, M. Mannan and P. De Filippi, "Now the Code Runs Itself: On-Chain and Off-Chain Governance of Blockchain Technologies", *SSRN Electronic Journal*, 2018. Available: 10.2139/ssrn.3340056.

[20] K. Aiyar, M. Halgamuge and A. Mohammad, "Probability Distribution Model to Analyze the Trade-off between Scalability and Security of Sharding-Based Blockchain Networks", *Ieeexplore.ieee.org*, 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9369563. [Accessed: 30- Sep- 2022]