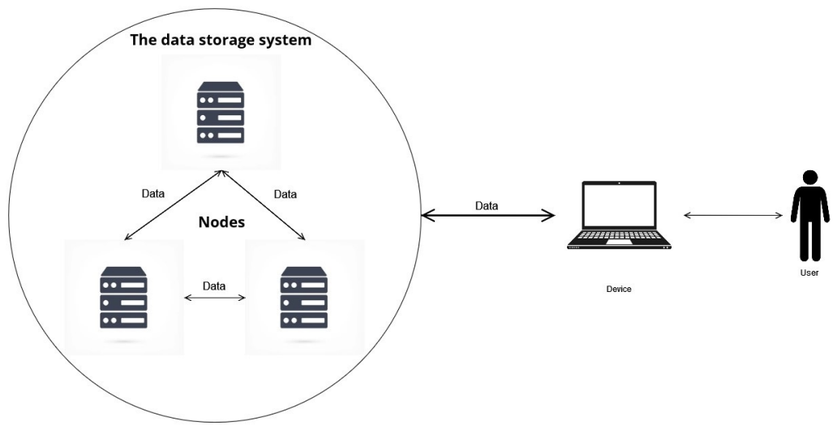
Comparison of Big Data Storage Solutions with the Hybrid Blockchain Architecture

ABSTRACT

A massive amount of medical and health data is produced for the diagnosis, monitoring, and treatment purposes. The gathered and stored content must be valid, complete, traceable, and immune to modification or deletion. The big data collection from the Internet of Things (IoT) and mobile devices emergence influenced data storage and processing solutions to be more reliable and decentralized. Wide usage of IoT technologies require secure data transitions and storage that could be ensured by private blockchain networks that create tamper-resistant records of shared transactions. Currently, there is a lack of comprehensive and specific comparison among widely recognized primary big data storage solutions such as SQL, NoSQL, or blockchain-based approaches. Considering data storage and the frequent issues of data loss that arise from it, the comparison becomes important on security aspects.

PROJECT DESCRIPTION

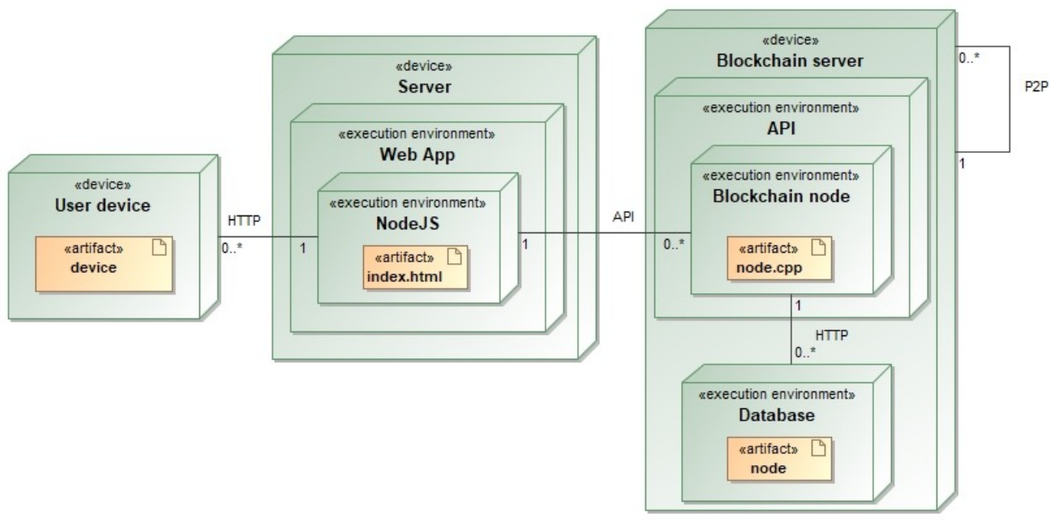
The aim of this project is to promote innovation in secure data storage through blockchain and traditional data storage solutions, so that users and practitioners involved in secure data storage are aware of the possibilities of secure data storage. This project combine traditional SQL and blockchain-based solutions while bringing together the principles of elliptic curve cryptography and multi-threading to securely and efficiently store user data.



PLATFORM ARCHITECTURE

There are two main parts to a data storage system:

- The **graphical user interface (on the server)** is designed to facilitate the viewing and manipulation of the data contained in the data storage system;
- An **application programming interface (API) on a blockchain server** that receives data from external interfaces and distributes it appropriately over a peer-to-peer (P2P) network by storing the data and performing the appropriate operations on its own and other nodes' data storage systems.



DATA STORAGE SOLUTIONS COMPARISON

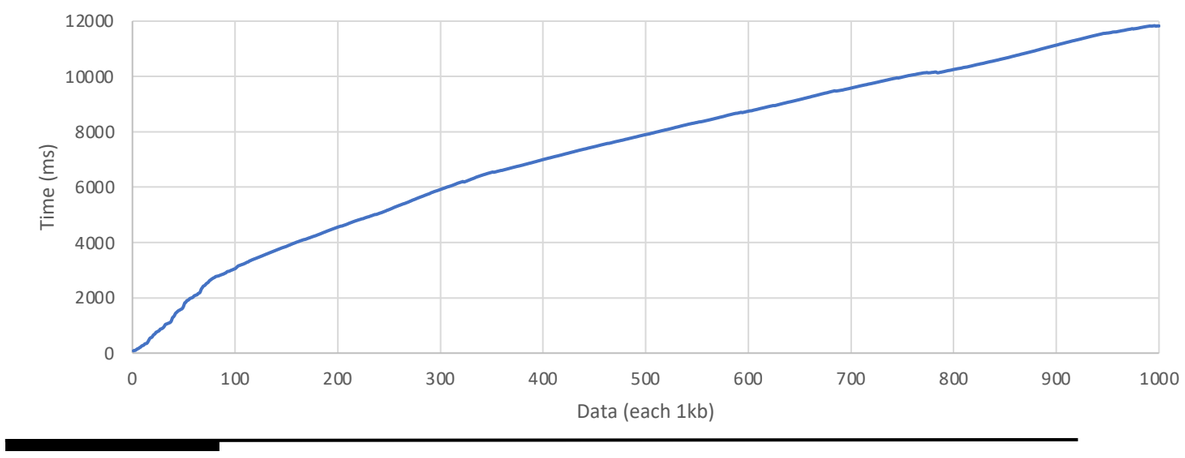
Traditional databases allow data to be modified and have a low response time, allowing many operations to be performed at the same time, unlike blockchains [1, 2]. Even though traditional databases may use an internal replication mechanism, they do not have the advantage of being able to replicate across many nodes, as in blockchains (only across a few) [3].

Criteria	Blockchain	SQL or NoSQL
Immutable	Yes	No
Operations	Insert/append	Create, read, update, delete
Number of nodes	Many	Few
Redundancy	Can be fully replicated within every node	Centralised, prone to single point of failure
Consensus	Majority of peers agree on outcome of transactions	Central authority
Latency	High	Low
Transaction cost	High	Low
History	In the ledger	Naturally not saved

RESULTS

According to the initial results, with three nodes and 1000 users (3000 users) sending 1kb of data (3000 kb) to each of them at the same time approximately 1000 (each 1kb) of data from one node is saved in 12000 ms (results represent averaged data when the same process was repeated 10 times). Averaging over three nodes and sending 1kb to each node at the same time results in 83ms, 10kb - 169.1ms, 100kb - 1667.61ms, 1000kb - 7438.974ms.

Data storage in a three-node database



CONCLUSIONS

This study proposes a hybrid blockchain architecture that combines traditional SQL and blockchain-based solutions to securely save users data with tamper-proof resistance. The proposed technique uses an unconventional consensus algorithm for faster agreement and data signing before saving.

A. Bendoraitis, T. Blažauskas, E. Butkevičiūtė

Faculty of Informatics, KTU

REFERENCES

[1] MCBEE, M.P. - WILCOX, C. Blockchain Technology: Principles and Applications in Medical Imaging. In Journal of Digital Imaging. 2020. Vol. 33, no. 3, p. 726.

[2] CASINO, F. et al. A systematic literature review of blockchain-based applications: Current status, classification and open issues. In Telematics and Informatics . 2019. Vol. 36, p. 55–81.

[3] MA, X. et al. Exploration of Radio Monitoring Data Storage Based on Blockchain Technology. In Proceedings - 2021 International Conference on Computer, Blockchain and Financial Development, CBFD 2021. 2021. p. 519–522.