Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 220 (2023) 832–837

Procedia Computer Science

www.elsevier.com/locate/procedia

The 4th International Workshop on Hospital 4.0 (Hospital)
March 15-17, 2023, Leuven, Belgium

Security and Immutability of Open Data in Healthcare

Tiago Guimarães[a*], Ricardo Duarte[a], João Cunha[a], Paulo Gomes[a], Manuel Filipe Santos[a]

[a]*Algoritmi Reasarch Center, School of Engineering, University of Minho, Azurém Campus, Guimarães, 4800-05, Portugal*

Abstract

Clinical data are sensitive data given the origin of the information. Since the implementation of health information systems, some issues such as interoperability, security, and privacy have been strongly questioned. Storing and consulting them raises the same concerns. Given these concerns, any attempt to introduce healthcare information systems must guarantee the security and privacy, integrity, and immutability of patient information. It is in this sense that blockchain technology and the openEHR open data model appear, as they manage to guarantee interoperability between systems, data security and guarantees about queries of each stored data.

In order to understand how to increase security and immutability in an implementation of open data models in hospitals, two distinct architectures were developed. In these architectures, several performance tests were carried out. To understand which of them represents more value to a health institution, an analysis of the results was prepared and, consequently, a discussion about them was held to be able to draw the respective conclusions.

Keywords Blockchain; Blockchain in Healthcare; Benchmarking; Hyperledger Fabric; OpenEHR

*Corresponding author. Tel.: +351-918-608-484.
E-mail address: tsg@dsi.uminho.pt

1. Introduction

The health sector is a sector that has unique requirements, and trusting the data from its activities is essential for its operations [1], [2]. With the growing amount of data generated, some problems arise, including unauthorized

1877-0509
10.1016/j.procs.2023.03.111

sharing, invasions, and theft of confidential data, in addition to practices such as the falsification of both medicines and patients. These propensities lead to people's suspect and doubts about the veracity of these institutions [2], [3]. To address these issues, it is important for these types of institutions to consider alternative approaches, and it is in this sense that blockchain technology emerges. A technology that offers a solution to the needs demanded by the sector, given its nature and characteristics [4], [5]. Allied to this technology comes the open data structure, openEHR. Where it enables secure and reliable structuring, management, storage, and switching of patient data across healthcare organizations. The main idea of this approach is to standardize concepts related to health used in databases or Electronic Health Record (EHR) systems in a set of libraries, called archetype [6].

The present work is divided into several sections, initiating with a brief introduction. In the section two is presented a literature review about OpenEHR and blockchain in healthcare. The next section presents the developed architectures, which will serve as the basis for the tests performed in section four. Finally, in sections five and six are present, the discussion of the results obtained and the final conclusions, respectively.

2. Background

Data security and integrity have always required special attention in a hospital environment. When information systems began to be implemented in these environments, ensuring their security was the highest priority given their sensitivity. The need to guarantee the confidentiality, integrity, security, privacy, and interoperability of this data is imperative [7]. To ensure these needs, the use of the openEHR open data model and blockchain technology arises.

In this sense, openEHR is an open data model that allows free access to health information specifications, used in the management, storage, and consultation of the electronic clinical file. The use of this model provides an interoperable framework that organizes clinical content with patient information, allowing integration with different health information systems. In terms of data security, it ensures their resilience, preserving them in a system from a historical and review perspective [8].

Whereas, blockchain technology emerges as a general-purpose technology with a presence and prominence in several areas, among which health stands out [5], [9]. This is an area where there are many opportunities for application of the concept, as there are a variety of problems that can be solved through the attributes offered by it [10], [11]. The use of this technology to support the exchange of health information can unleash the true value of an interoperable system, offering a distributed framework capable of guaranteeing access to patient information securely.

Information stored on the blockchain is universally available to specific individuals through public and private key mechanisms, thus allowing patients to share their information with institutions in an easier and more secure way. Deploying a transaction layer on the blockchain can help create a collaborative and trusting ecosystem for sharing information to create insights to improve healthcare system efficiency [12].

3. Architectural Solution

By integrating Blockchain technology with the open data model, openEHR, it is possible to maintain data interoperability and standardization of all EHRs, in addition to maintaining data veracity, privacy, and security.

For the development of our case study, two scenarios were considered. Scenario 1 represented in Figure 1 and scenario 2 represented in Figure 2.

Analysing the flow of scenario 1, shown in the following figure, it starts with the insertion of information and data by the patients. These are transformed into EHRs and sequentially processed and analysed according to the specifications, modelled in the open data model, openEHR, in the APIS layer, through the gateway. That said, Hyperledger Caliper was used to perform a series of insertions in the blockchain. Next, the blockchain will store all the transitions for them to be consulted in a secure, immutable, and private way. Finally, Hyperledger Caliper evaluates transitions in terms of success, speed, maximum, minimum, and average time to send and receive a response, and the average number of transactions processed per second. The mentioned metrics are represented in an HTML page.

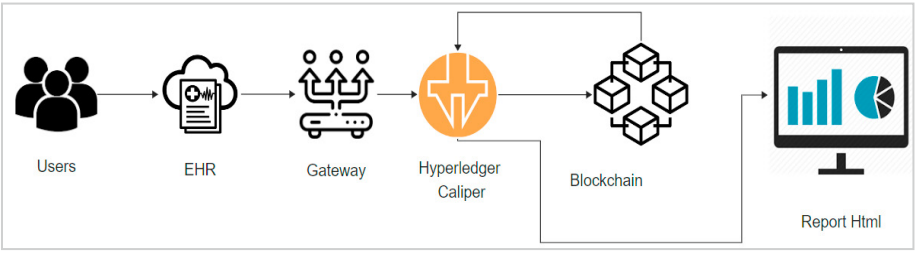

Figure 1 – First Scenario Architecture

Starting with the analysis of scenario 2, shown in the following figure, the flow starts in the same way as in the scenario presented above. This begins with the insertion of clinical information of the patients in the systems present in the institution. These are transformed into EHRs and sequentially processed and analysed according to the specifications, modelled in the open data model, openEHR, in the APIS layer, through the gateway. Following the flow, it is possible to observe a fork. Following the flow of the upper arrow, the data will be stored in the hospital's database, to be further processed and transformed according to the needs of the hospital's stakeholders. Next, in the lower path, a hash block is created for each object, through the encoding process. Associated with the block is an object ID to check which person we are referring to. The main objective of this technique is to validate if there has been any intrusion or alteration to the data, providing greater security. If the objective is to change some type of data in the object, the md5 will be changed and will be stored on the blockchain again. Going with the flow, the Hyperledger Caliper triggers people's entries into the blockchain. However, in this case, the constitution of the object changes and becomes just an ID and the hash. Next, the blockchain will store the transactions and will play an important role in verifying whether a hash differs from the one recorded. Finally, the Hyperledger Caliper uses metrics to measure performance, which are presented in an HTML page.

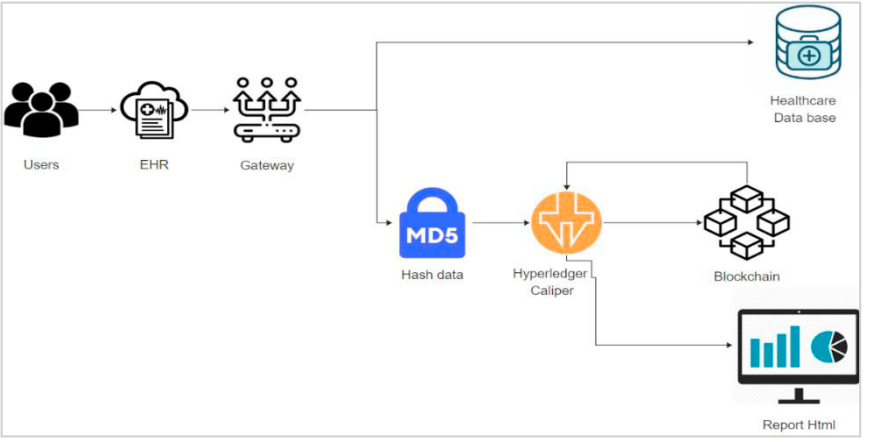

Figure 2 - Second Scenario Architecture

4. Tests Performed

In this chapter, a visual demonstration of the performance that the network obtained in the various stress tests provided to it will be carried out. Thus, it will be possible to measure and understand whether the objective of increasing the security and immutability of an implementation of open data models in a hospital environment has been achieved. In this income statement, two tests were performed, one for the first scenario and another for the second scenario. At each test two types of graphs are shown, one evaluates the performance of the network in the gradual submission of people. The other evaluates the amount of memory that the network used in total.

4.1. Test 1 – Scenario 1

For this test, four insertions of 20000 people were performed on the blockchain. Initially, the container was restarted. As the insertions were performed, the processing capacity of the network was decreasing, gradually increasing the average latency. This is explained by the increasing amount of data stored in the blockchain. It is noteworthy that the processing speed was 5 transaction per second (TPS), adding this to the physical capacity of the machine used for testing, the processing time increased considerably. A positive aspect that goes against the speed stability corresponds to the absence of failures.

In the following figure, it is possible to see a growing increase in the maximum and average latency variables and a slight variation in the minimum. As mentioned, the amount of data entered is twenty thousand in each of the four iterations, thus pushing the capacity of the blockchain to the limit. The previous statement is proved due to the high value of both the average and maximum latency from the second to the fourth iteration performed. Going from 583.74 seconds to 1035.24 seconds and from 834.89 seconds to 1935.76 seconds respectively.

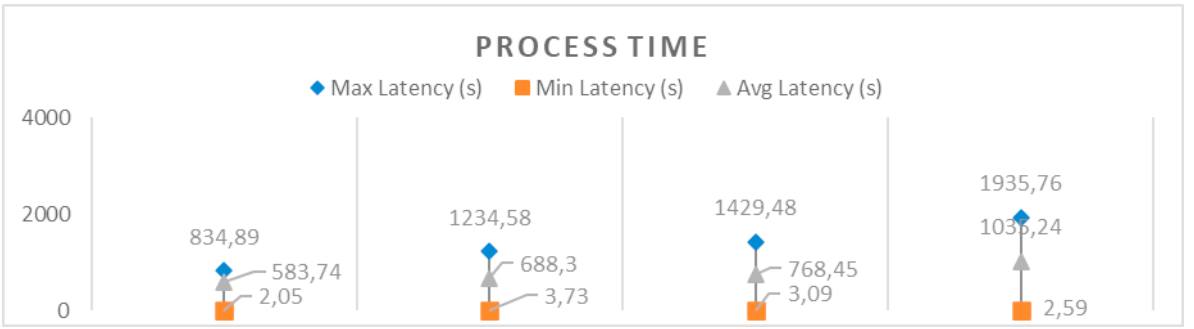

Figure 3 – Test 1 – Scenario 1: Process Time

Regarding memory, the following figure shows the total memory in megabytes (MB) of each container in the respective people submissions. The memory evolution progressively increased from 1679.55 MB in the first case to 2566.96 MB in the last case, demonstrating a constant increase. With each iteration, there was an increase between 200 – 300 MB. In this way, it is possible to conclude that as the volume of data increases, the memory used by the system also increases, which will cause a slower system.

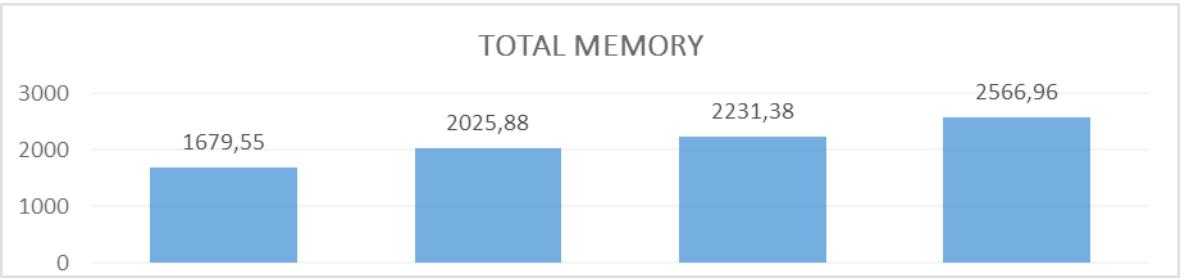

Figure 4 – Test 1 – Scenario 1: Memory Usage

4.2. Test 1 – Scenario 2

For this test, four insertions of 20000 people were performed on the blockchain. As the iterations were carried out, the network performance naturally decreased, which caused an increase in processing time. From the first to the last iteration, the average and maximum latency had a significant increase, going from 60.33 seconds to 763.4 seconds and from 179.35 to 1296.31, respectively, as we can see in the following figure. In this way, it caused a significant increase in processing. The problem inherent in this test is based on the limit of the send rate to 5 seconds per transaction, causing a progressive increase in the time of sending data and respective receiving response.

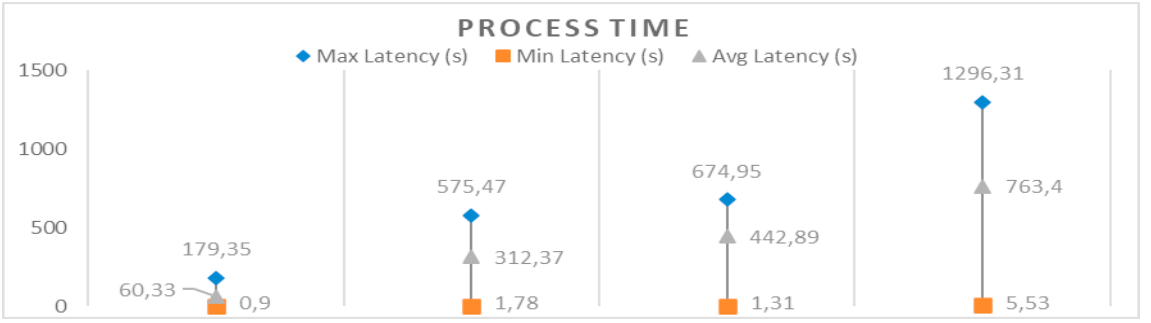

Figure 5 – Test 1 – Scenario 3: Process Time

Regarding the memory tests, the following figure shows the total memory in megabytes (MB) of each container in the respective people submissions. The memory evolution progressively increased from 529.39 MB in the first iteration to 1631.93 MB in the last iteration, demonstrating considerable but consistent evolution.

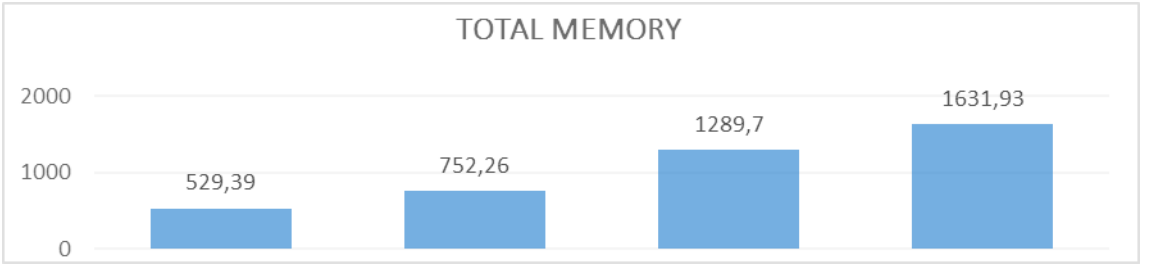

Figure 6 – Test 1 – Scenario 4: Memory Usage

5. Discussion

Both tests performed are similar, in each one of them was performed the insertion of 20000 people in 4 different iterations with a sending rate of 5 TPS. However, it is important to note that the chaincode used differs, chaincode person and chaincode personSecurity respectively.

For the test performed in the first scenario, the percentage change rate of the average processed time performance is approximately 77.34%. Starting from the value of the first iteration of 583.74 seconds, its performance increased the average processing time by 77.34% reaching a value of 1035.24 seconds in the last iteration. On the other hand, in the test performed in the second scenario, the percentage change rate of the average processed time performance is approximately 1165.37%. Starting with the first iteration with an average time of 60.33 seconds, there was an increase of 1165.37% in processing time to reach the fourth iteration, reaching a value of 763.4 seconds.

That said, when comparing the percentage change rates of the two tests carried out, it is possible to see that there is more instability in the test carried out in scenario 2, where the personSecurity chaincode is used when compared to the test carried out in scenario 1 where the person chaincode is used. Where it presents in all iterations a constant average value over the iterations, revealing a better scalability for this case.

The other metric analysed is memory usage. In the case of scenario 1, the memory remained constant and throughout the iterations it did not grow much, going from 1679.55 MB in the first iteration to 2566.96 MB in the fourth iteration, representing an increase of approximately 53%. The scenario 2, on the other hand, increased considerably from the second iteration, presenting an increase of approximately 117% from the second iteration (752.26 MB) to the fourth (1631.93 MB). Scenario 2 should not have such a considerable increase in memory since the composition of its object does not occupy as much space as that of scenario 1.

Finally, we highlight scenario 1 as the best scenario for the institution. This scenario as it deals with a more complex and composite object is the one chosen for the performance level. The above does not show many variations in processing time or memory. While the second scenario, being a simpler object, should have more competitive processing times.

6. Conclusion

To design the scenarios described, a review of the literature was necessary to understand what currently exists and from there to determine the best way forward for the development of the private blockchain network. Hyperledger Fabric was the framework chosen for this purpose, as it is an open-source tool that meets all the needs required for building a blockchain network.

For the case study, two scenarios were developed, and a test was carried out in each of them. Where, in both cases, 4 insertions of 20000 people were performed in each of the iterations. After analysing the results and discussing them, it is concluded that scenario 1 is the best scenario. In terms of performance, it does not present large variations in processing time or memory. In addition, sending data to the blockchain would not exceed 5 TPS.

A future work to be considered involves creating a test, in each of the scenarios, where there is a configuration that allows the system to be free to select the speed at which it can send the data. Subsequently, it would be necessary to make comparisons again to verify which would be the best scenario.

References

[1] M. M. H. Onik, S. Aich, J. Yang, C.-S. Kim, and H.-C. Kim, *Blockchain in Healthcare: Challenges and Solutions*. Elsevier Inc., 2019. doi: 10.1016/b978-0-12-818146-1.00008-8.

[2] A. Hasselgren, K. Kralevska, D. Gligoroski, S. A. Pedersen, and A. Faxvaag, "Blockchain in healthcare and health sciences—A scoping review," *Int. J. Med. Inform.*, vol. 134, no. May 2019, p. 104040, 2020, doi: 10.1016/j.ijmedinf.2019.104040.

[3] A. Gamal, S. Barakat, and A. Rezk, "Standardized electronic health record data modeling and persistence: A comparative review," *J. Biomed. Inform.*, vol. 114, no. December 2020, p. 103670, 2021, doi: 10.1016/j.jbi.2020.103670.

[4] M. Hölbl, M. Kompara, A. Kamišalić, and L. N. Zlatolas, "A systematic review of the use of blockchain in healthcare," *Symmetry (Basel).*, vol. 10, no. 10, 2018, doi: 10.3390/sym10100470.

[5] M. Prokofieva and S. J. Miah, "Blockchain in healthcare," *Australas. J. Inf. Syst.*, vol. 23, pp. 1–22, 2019, doi: 10.3127/ajis.v23i0.2203.

[6] T. Ribeiro, S. Oliveira, C. Portela, and M. Santos, "Clinical workflows based on OpenEHR using BPM," *ICT4AWE 2019 - Proc. 5th Int. Conf. Inf. Commun. Technol. Ageing Well e-Health*, no. Ict4awe, pp. 352–358, 2019, doi: 10.5220/0007878203520358.

[7] E. Smith and J. H. P. Eloff, "Security in health-care information systems - Current trends," *Int. J. Med. Inform.*, vol. 54, no. 1, pp. 39–54, 1999, doi: 10.1016/S1386-5056(98)00168-3.

[8] D. Oliveira *et al.*, "Management of a pandemic based on an openEHR approach," *Procedia Comput. Sci.*, vol. 177, pp. 522–527, 2020, doi: 10.1016/j.procs.2020.10.072.

[9] C. C. Agbo, Q. H. Mahmoud, and J. M. Eklund, "Blockchain technology in healthcare: A systematic review," *Healthc.*, vol. 7, no. 2, 2019, doi: 10.3390/healthcare7020056.

[10] L. Bell, W. J. Buchanan, J. Cameron, and O. Lo, "Applications of Blockchain Within Healthcare," *Blockchain Healthc. Today*, vol. 1, pp. 1–7, 2018, doi: 10.30953/bhty.v1.8.

[11] C. C. Agbo and Q. H. Mahmoud, "Blockchain in healthcare opportunities, challenges, and possible solutions," *Int. J. Healthc. Inf. Syst. Informatics*, vol. 15, no. 3, pp. 82–97, 2020, doi: 10.4018/IJHISI.2020070105.

[12] J. Lopes and J. L. Pereira, "Blockchain technologies: Opportunities in healthcare," *Adv. Intell. Syst. Comput.*, vol. 850, pp. 435–442, 2019, doi: 10.1007/978-3-030-02351-5_49.