LEBANESE AMERICAN UNIVERSITY

A Secure and Scalable Sidechain Model for Fog Computing in Healthcare Systems

By

Ali Amhaz

A thesis Submitted in partial fulfillment of the requirements for the degree of Master of Science in Computer Science

School of Arts and Sciences

August 2022

LAU
الجامعة اللبنانية الأميركية
Lebanese American University

THESIS APPROVAL FORM

Student Name: Ali Amhaz I.D. #: 201604732

Thesis Title: A Secure and Scalable Sidechain Model for Fog Computing in Healthcare Systems

Program: Master of Science in Computer Science

Department: Computer Science and Mathematics

School: Arts and Sciences

The undersigned certify that they have examined the final electronic copy of this thesis and approved it in Partial Fulfillment of the requirements for the degree of:

Master of Science in the major of Computer Science

Thesis Advisor's Name: Ramzi A. Haraty

Signature: [redacted] Date: 22 / 08 / 2022 (Day / Month / Year)

Committee Member's Name: Samer Habre

Signature: [redacted] Date: 22 / 08 / 2022 (Day / Month / Year)

Committee Member's Name: Sanaa Kaddoura

Signature: [redacted] Date: 22 / 08 / 2022 (Day / Month / Year)

LAU
الجامعة اللبنانية الأميركية
Lebanese American University

THESIS COPYRIGHT RELEASE FORM

LEBANESE AMERICAN UNIVERSITY NON-EXCLUSIVE DISTRIBUTION LICENSE

By signing and submitting this license, you (the author(s) or copyright owner) grants the Lebanese American University (LAU) the non-exclusive right to reproduce, translate (as defined below), and/or distribute your submission (including the abstract) worldwide in print and electronic formats and in any medium, including but not limited to audio or video. You agree that LAU may, without changing the content, translate the submission to any medium or format for the purpose of preservation. You also agree that LAU may keep more than one copy of this submission for purposes of security, backup and preservation. You represent that the submission is your original work, and that you have the right to grant the rights contained in this license. You also represent that your submission does not, to the best of your knowledge, infringe upon anyone's copyright. If the submission contains material for which you do not hold copyright, you represent that you have obtained the unrestricted permission of the copyright owner to grant LAU the rights required by this license, and that such third-party owned material is clearly identified and acknowledged within the text or content of the submission. IF THE SUBMISSION IS BASED UPON WORK THAT HAS BEEN SPONSORED OR SUPPORTED BY AN AGENCY OR ORGANIZATION OTHER THAN LAU, YOU REPRESENT THAT YOU HAVE FULFILLED ANY RIGHT OF REVIEW OR OTHER OBLIGATIONS REQUIRED BY SUCH CONTRACT OR AGREEMENT. LAU will clearly identify your name(s) as the author(s) or owner(s) of the submission, and will not make any alteration, other than as allowed by this license, to your submission.

Name: Ali Amhaz [redacted]

Signature: [redacted] Date: 22 / 08 / 2022
Day / Month / Year

LAU
الجامعة اللبنانية الأميركية
Lebanese American University

PLAGIARISM POLICY COMPLIANCE STATEMENT

I certify that:

1. I have read and understood LAU's Plagiarism Policy.
2. I understand that failure to comply with this Policy can lead to academic and disciplinary actions against me.
3. This work is substantially my own, and to the extent that any part of this work is not my own I have indicated that by acknowledging its sources.

Name: Ali Am [redacted]

Signature: [redacted] Date: 22 / 08 / 2022
Day / Month / Year

DEDICATION

I will dedicate this thesis to my family, that supported me during its accomplishment.

ACKNOWLEDGMENT

This work wouldn't have been possible without the support of my advisor Dr. Ramzi Haraty who gave me the necessary instructions and feedback. Also, special thanks to my committee Dr. Samer Haber and Dr. Sanaa Kaddoura, who were very cooperative.

A Secure and Scalable Sidechain Model for Fog Computing in Healthcare Systems

Ali Amhaz

ABSTRACT

With the enormous amount of data produced daily, cloud and fog computing presented efficient and effective models for real-time data exchange. Nevertheless, this technology came with a cost at the security level, where it became an easy target for malicious actions that could directly spread throughout the model. Blockchain, a recent and promising technology, was a suitable solution for securing the transactions in the fog environment because of the distributed ledger structure that makes it resistant to many attacks. Scalability, however, introduced the main drawback for a blockchain by making it inefficient in some real-world applications, especially in the medical field, which includes a lot of data exchange. This work will suggest a scalable and secure model for fog and cloud computing in healthcare systems that depend on sidechains and the clustering of the available fog nodes. The importance of the model is highlighted, and experimental results showed promising outcomes.

Keywords: Blockchain, Sidechain, Fog Computing, Cloud Computing, Scalability, Healthcare.

Table of Contents

List of Tables

LIST OF FIGURES

Chapter 1

Introduction

The technological revolution in the last century has led to a significant development in the software and hardware of information systems. For example, in the old banking systems, committing a transaction needed the manual assistance of one or more employees. However, presently, all the banking services are automated and allow clients to do transactions from their homes or any shop using online applications and micro hardware (i.e., electronic ships).

A significant part of this enormous development was the proliferation of the Internet of Things (IoT) devices. Those devices, which support connecting to the internet network, granted a variety of data manipulation actions such as gathering, transmitting, and processing. In addition, they helped in dispensing many human-controlled activities that consume time and resources. Many fields started using IoT devices because of their low prices and the ability to perform critical tasks without human supervision. Healthcare institutions, including hospitals, started using such devices to keep track of patient's health records (i.e., blood pressure, temperature) and add them to the central system for later use. Moreover, modern agriculture adopted IoT sensors connected to the internet to monitor the soil state for a better harvest.

This enormous development of information systems and the amount of data produced (especially by IoT devices) brought the need to invent and enhance those systems to satisfy the growing demands like storage, processing power, and availability. Cloud computing came to

solve these problems, offering several services to facilitate data manipulation. It allowed the accomplishment of many tasks using remote servers provided by several international companies (i.e., Google, Apple). For example, cloud storage offered by Google supplied the users with terabytes of storage at a low cost. Moreover, it eliminated the risk of losing the physical data resulting from any emergency. In addition, cloud services enhance the deployment of large programs that needs vast computer resources and human management.

Although cloud computing presented the solution to many problems, it raised others. Such a technology consists of a centralized structure that serves millions of users in the same place; thus, causing unwanted latency in some critical applications. For example, in automated car projects, the response time and availability are crucial measures that can lead to life-threatening problems. Those cars need quick operations toward any action that could happen on the roads (i.e., a child crossing the street). Based on that, fog computing [1] came as a solution to give a better performance. It provided services similar to cloud computing but with better performance. The distributed and close-to-user structure helped a lot in increasing the response time with much fewer failures.

Fog systems were secure compared to cloud computing because of the distributed architecture. Nevertheless, the communication between the different fog nodes represented the main vulnerability that a hacker could exploit. For instance, a malicious transaction targeting a specific fog node could spread throughout the system, making it very hard to recover to the original state. Moreover, the level of damage would be more in the case of heterogeneous models [2]. And to tackle this problem, researchers proposed two different

approaches [3, 4]. The first approach depends on preventing malicious transactions before entering the system, while the other suggests detecting and recovering the damage.

In the case of detecting and recovering malicious actions, studies focused on building efficient and effective algorithms [2, 5] that could scan the system for unusual behavior and directly start the recovering process in the case of attacks. For instance, in [6], the researchers focused on machine learning as a way to discover any intrusion and try to recover it. On the other hand, the prevention systems mainly focused on blockchain as a technology to approve any transaction before entering the fog system. For example, in [7], the authors presented a blockchain model to protect the system and facilitate data exchange for doctors.

Recently, blockchain technology became a target for many applications because of its high security and the ability to control the flow of transactions to any system, especially in fog networks. This combination (blockchain + fog) helped to filter the transactions of IoT devices by forcing the proof of work and validation between different nodes [8]. Blockchain effectiveness and efficiency are measured using a set of metrics to study its scalability and compatibility with the given systems [9].

1.1 Overview

This section will present an overview of the fundamental subjects tackled in this thesis.

1.1.1 IoT Devices

IoT devices are connected to the internet for gathering, transmitting, or processing data. IoT devices are deployed in many fields (see Figure 1) because of their low prices and the ability to manipulate data in severe environments.

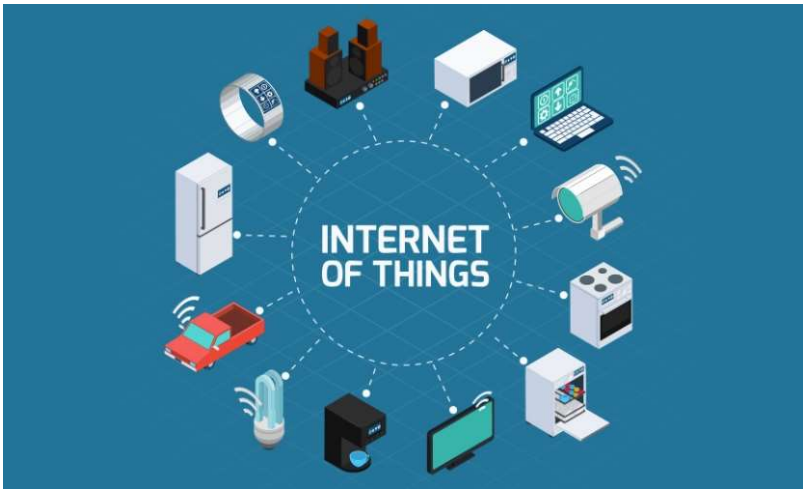

Figure 1: IoT Devices

Concerning healthcare systems, IoT devices found their place as a way to monitor patients' medical measures and provide instantaneous alerts in the case of emergencies. Moreover, they are facilitating the process of giving medicine to the clients. For instance, a recent IoT invention had the role of tracking the glucose level in blood and injecting the correct amount of insulin [10].

The benefits of IoT could be illustrated by:

Data Gathering: IoT devices facilitated the data collection process by their ability to work in severe environments which are hard to be reached by humans. For example, IoT devices play a critical role in managing space missions and reducing human interactions [11].

Efficiency: Because of their self-supervised nature and their ability to work under complex conditions, IoT devices are considered an efficient and effective way for data manipulation. In addition, they can provide more working hours in comparison to some electronic equipment that cannot work without human interaction.

Reduce human supervision: These devices are capable of monitoring their functions without human control because of their connectivity to the internet and the algorithms that specify their work.

Less price: In comparison with human-controlled equipment, these devices offered many services at low manufacturing costs. Moreover, they eliminated the cost of human supervision and required less maintenance.

Although IoT devices bring many benefits to information systems, they still suffer from drawbacks:

Complexity: The devices with their self-supervised structure are more complex to be embedded and integrated into any environment. They require planning and designing to get the required performance.

Security: The low processing power and the high dependency on the internet make the devices an easy target for many types of attacks. For example, a small-scale denial of service attack could stop its functionality because of the limited ability to handle user requests.

Network dependency: As it is well known, IoT devices perform their tasks through the internet connection; thus, any network failure will interrupt their activities and lead to delays.

Privacy: The IoT devices collect data without human supervision; for that reason, they could violate the privacy of the surrounding environment by sending unauthorized data.

1.1.2 Cloud Computing

Cloud computing is a set of facilities that allow remote access to storage, processing power, networks, etc. It is provided over the internet network to get better services and quick access. Its main advantages are:

Security: Nowadays, ensuring the security of the data and clients is one of the most critical factors for a successful information system. For that reason, cloud companies ensure the safe storing of the data uploaded by the customers. For instance, Google drive [12] encrypts the data stored on its side using a specialized encryption algorithm. Moreover, such companies authorize the users to check their identities and if they have the right to access the data. And with the evolution of such technology, researchers suggested new approaches to protect the cloud, especially using blockchain technology which guarantees a controlled data manipulation process [13].

Price: Compared to the prices of computer hardware needed to accomplish large computing tasks, cloud services presented an innovative and low-cost solution to get sufficient processing power. Such a service can save the expenses of establishing the information system on the client-side. Those expenses include the money needed to buy the hardware, set up the environment, maintenance, and power price. For instance, in countries with expensive power prices, cloud computing could present the best way to make a profit.

Effectivity: With the enormous processing power offered by cloud computing, many clients benefit from the ability to run large programs without being limited to the power of the available hardware. Cloud companies offer several services to satisfy the growing needs of users. For example, the clients could rent servers by selecting the needed specifications (i.e., memory, number of cores, bandwidth).

Availability: Cloud companies spend millions of dollars on providing a reliable service that can satisfy the client's needs. They establish different backup locations in the case of disasters to ensure the safety of the data. In addition, they ensure that the data is available upon request.

Speed: Speed is an important factor when dealing with cloud computing. Clients could store an enormous amount of data that should be quickly and smoothly accessible upon their request. Cloud companies ensure a satisfactory download/upload rate to guarantee a good experience when using the data. In addition, they provide a scheduling algorithm to assure fairness in accessing the available resources.

However, cloud computing does have some limitations:

Security: Although cloud computing could provide security on some sides, giving the data by itself to a third-party entity could be a problem. For instance, in 2014, dropbox deleted by mistake an enormous amount of users' data [14]. In addition, the central architecture of the cloud makes it a target for Denial of Service attacks (DoS). Attackers could focus on a specific server causing damage to millions of users.

Flexibility: Because of the global centralized structure and company-based cloud services, customizing the data storage and processing power becomes a rigid part. Customers need to deliver their preferences to the cloud agencies; thus, losing time and money.

Performance: Cloud architecture is usually central, serving millions of users from a specific location. For that reason, cloud services could not cover the demands, especially in the applications that require fast response time. For example, automated car projects, which are very popular nowadays, require a quick response to the changing environment (i.e., unexpected accidents).

Network Dependence: Every cloud service relies on the internet network. For that reason, good quality internet represents a critical factor for smooth use. Problems in connection will directly

reflect on the performance of the users. For instance, natural disasters can corrupt the internet network, and any work which depends on cloud computing would stop.

Changing Providers: As it is well-known, there is no single provider for cloud services. For that reason, moving from one provider to the other can lead to difficulties because of the different architecture and systems.

1.1.3 Fog Computing

Like cloud computing in terms of services, fog computing is a more decentralized system that facilitates storing, processing, and managing data. It is located in a middle way between the cloud services and the clients. Moreover, it is considered a complementary part of the cloud. Fog computing offers some significant features, including:

Scalability: As previously mentioned, Fog computing supports a decentralized structure that focuses on distributing resources over several sights. For that reason, fog services will eliminate the overhead of having a centralized structure that serves millions of users at the same time. Moreover, such an architecture will filter the data before uploading it to the cloud; thus, reducing the pressure on the cloud resources. For example, the street cameras, which monitor the traffic, will benefit from the fog network by uploading only (to the cloud) the critical events instead of saving the whole view, which wastes the system resources.

Security: Fog computing, with its separated nodes, presents default protection against threats. The structure will make it harder for the attacker to locate the useful data. In addition, this distribution in nodes makes it easier to identify any potential threat before happening.

Response Time and Latency: While performing computing operations, the fog nodes are technically near the users; hence, they will benefit from a better response time and less

latency. These measures are critical in some real-life applications where timing is more important than the data itself. For instance, in the medical field, the speed in executing an action could save the life of a human. Another example is the fire-alarm systems, where performing the right action at the right time will help in avoiding losses.

While Fog computing is desirable in today's computing paradigms, it faces a few challenges:

Complexity: Dividing the processing power over the fog nodes is beneficial in terms of performance but a problem in its complexity. The deployed programs should adapt to the decentralized structure; thus, adding more complexity to the development process. For instance, in some programs, data should be divided before processing and regathered after that.

Security: Although fog computing presents security by its architecture, malicious transactions could spread throughout the system, causing damage. Attackers can take advantage of the decentralized systems that exchange and process data between its nodes. For instance, if a malicious database query is dependent on data from another fog node, both nodes will get affected. In addition, recovering from this damage is complex and costly.

Cost: Establishing a fog system requires a lot of implementation and hardware. For that reason, fog computing structures are among the most expensive systems to build. In addition, they have high power consumption and need maintenance.

1.1.4 Blockchain Technology

In 2008, Satoshi Nakamoto [15] introduced blockchain technology through a paper that explained the implementation of Bitcoin. This new technology represented a revolution in presenting a new security mechanism to protect and manage the data. It relied on a

decentralized structure that did not require a single point of control. Blockchain, with its new architecture, was self-supervised by its users and by adopting smart contracts.

Blockchain is a distributed ledger (see Figure 2) that records all the transactions done by the users. Each user has a copy of the chain, and it is updated regularly through time.

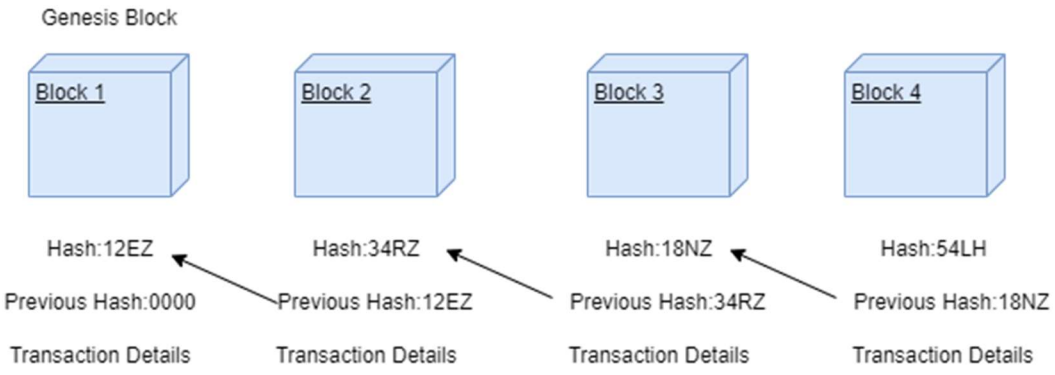

Figure 2: BlockChain Architecture

Every block consists of a "hash" that reflects the content of the chain and protects it from being altered. Moreover, the block has a pointer that records the hash of the previous one and the transaction details. The first block is always called the "Genesis".

Every time a user wants to commit a transaction, a new block is established with all the details. The new block needs the approval of the majority/all the users. This mechanism is critical to monitoring the flow of transactions and preventing hackers from inserting unauthorized transactions. Once the block is created, it will point to the hash of the previous block to prevent any alteration to the original chain. Finally, no one can modify the transactions once it is added to the ledger.

The security mechanism in the blockchain relies on 3 main points:

- By recording the hash of the previous block, the hacker is obliged to change the hash of all the preceding blocks in the system.
- The distributed ledger structure forces the attacker to gain control over the majority of the network to alter or insert blocks, which is impossible.
- The voting mechanism also protects the newly created blocks and ensures they follow the system's policy.

Blockchain technology entered many advantages to information systems:

Security: The distributed ledger structure and the proof of work protocol make it impossible for attackers to alter the chain. The attackers need to take control over the majority of the user, which is an impossible process.

Privacy: The blockchain's users could hide their identity, and no one could track the origin of the transactions. The encryption and decryption strategy adopted in this technology allows the users to own their data and prevent any unauthorized. Moreover, the ledger structure makes it hard to track the important information of the transaction like time, location, and users.

Trust: Being moderated by all the users, blockchain represented a trustworthy environment to commit transactions and exchange assets. Blockchain allows the users to safely exchange assets through the ledger without any risk because of the smart contracts that initiate trust between the different entities of the transactions.

Simplicity: Blockchain is very simple to use and does not require professional experience. Creating online wallets that allow the storage of the chain's currency is easy and doesn't require any conditions. In addition, using the chains is not restricted by location, country, nationality, etc.

Despite their advantages, blockchain has some disadvantages like:

Scalability: Because millions of users are using blockchain technology, the proof of work will take more time to be accomplished. For that reason, scalability represents a significant problem that could affect the performance of the whole system.

Illegal Use: The anonymity of the users made blockchain an attractive environment for criminals around the world. Many reports stated that more than 76 billion dollars were for illegal transactions on the Bitcoin platform in 2019 [16].

1.1.5 Sidechains

Sidechains represented a practical solution to overcome the problem of scalability. This new solution suggests the implementation of small chains that could exchange assets with the main chain when needed. It proved its importance by decreasing the time to add new blocks and by creating local privacy in the smaller chains.

Side chains could have a different structure and mechanisms than the main chain. Such technology requires specific protocols that ensure a smooth transfer of assets between the two communicating ledgers [17].

Two-way Peg Protocol

It is the most critical protocol to ensure the data exchange between two chains. It ensures the integrity of the data transferred by implementing smart contracts between two entities.

This protocol can be of two types symmetric and asymmetric. The symmetric protocol deals with chains with similar properties and structures. On the other hand, the asymmetric ensures the cross-chain communication between different chains (i.e., Bitcoin and Ethereum).

Sidechains entered a lot of advantages to the blockchain technology like

Scalability: Because the sidechain removes the bottleneck effect of having a central chain. It represents a suitable solution for the scalability problem by dividing the data over a set of smaller ledgers with a smaller number of participants. This small number of users per chain reduces the time required for validating the new blocks; thus, leading to better throughput and performance.

Local Privacy: The sidechains are separated from the main chains in terms of block creation and data manipulation. For that reason, the transactions that are done within a sidechain are local and provide privacy for its users.

Protection for the Main Chain: Sidechains allow testing new features and deploying policies without affecting and altering the main chain. For that reason, they protect the root chain from bugs and other problems.

Despite its benefits, sidechains still suffer from problems like

Complexity: Cross-chain protocols (like 2-way-peg) are not easy to implement and require attention, especially in the case of asymmetric chains. In fact, with the presence of new chains and features, cross-chain algorithms are becoming more complex and harder to develop.

Price: Sidechains require more human resources in the development process. For that reason, it is considered expensive in comparison with the regular blockchain. Moreover, they involve more maintenance expenses.

The paper [18] was the first to mention side chains as a potential solution to overcome the scalability problem in blockchain technology. It emphasized its importance in establishing private chains owned by specific institutions. In addition, they highlighted that some of its

problems are the complexity and the difficulty in communication between two chains with different structures.

1.2 Problem Statement

The fog architecture, with its advantages, became a target for many information systems, including healthcare [2]. However, such a structure had a security problem by being vulnerable to malicious transactions that could spread throughout the whole system. This infection is very hard to recover due to the decentralized structure of the fog system and the exchange of data between the different nodes. For example, a database query in a node could depend on the affected data from another; thus, transmitting the bad data.

The security solution for the fog architecture is divided into two categories. The first is detection and recovery, where the researchers tried to propose algorithms and models that could effectively detect and recover the effect of malicious transactions [2, 5]. Others suggested prevention mechanisms to avoid those transactions before happening [19, 20].

Many researchers [13, 19, 21] suggested the use of blockchain technology as a way to overcome this problem. This technology could secure the transactions by the validation process that could stop any suspicious actions committed in the fog network. In addition, the distributed ledger structure could provide a detailed record of queries performed by the different nodes.

Although blockchain technology combined with the fog architecture represented a secure and trustworthy information system, they suffered from the problem of scalability with the increase in the number of fog nodes [8]. This problem would significantly decrease the

performance of the whole system, especially in the case of HealthCare which includes millions of transactions every day.

Based on all the above, this thesis will focus on presenting a secure and scalable blockchain approach for fog computing in the healthcare system. This approach will focus on the ability to implement several sidechains; thus, decreasing the time needed for adding transactions.

1.3 Contribution of the Thesis

Despite the superiority of fog computing over the cloud, it still suffers from a set of security vulnerabilities that are hard to handle. Those vulnerabilities came from the decentralized architecture of fog that makes it easy to spread malicious transactions between the correlated nodes. Moreover, the separated structure imposes many challenges in terms of tracking the attacks and recovering their effects.

BlockChain, a recent and promising technology, protected the environments that include committing transactions and exchanging data. Based on that, this thesis will suggest a blockchain model that secures the fog nodes using a sidechain approach. The model will cluster the nodes based on their frequency of communication to form a sidechain and ensure local privacy between them. Within the same cluster, when a node needs to commit a transaction, it is approved first by the majority of nodes, then a new block is created on the ledger. On the other hand, when there is communication between nodes from different clusters, the transactions are committed through a mainchain that is formed from the unclustered nodes.

Based on that, the contribution in this thesis will be:

- Blockchain technology is used to protect the data shared between different fog nodes.

- The model avoids scalability problems by establishing sidechains in the clusters of the fog nodes.
- The suggested approach eliminates the need for centralized communication between fog nodes in different clusters.
- The model ensures privacy within the clusters.

1.4 Organization of the Thesis

In Chapter 2, the thesis will go over the literature review related to the topic. Then, Chapter 3 will suggest a new model that provides a secure and scalable blockchain solution for a healthcare system. Chapter 4 will show and analyze the obtained results after the simulations. Finally, the conclusion will be presented in Chapter 5.

Chapter 2

Related Work

In this chapter, the thesis will tackle the most relevant studies that focus on using cloud and fog technologies in information systems. In addition, it will go over the different security measures, especially the blockchain and sidechain models, for better data exchange in the fog environment. Those measures are divided into two categories: detection and recovery methods and prevention methods. In the first one, the authors will try to assess the damage of the attacks and provide the necessary recovery algorithms. On the other hand, the prevention solutions focus mainly on blockchain technology to assure trustable and approved transactions.

2.1 Cloud and Fog Computing in Information Systems

Since the existence of cloud computing, many system models began adopting its architecture because of the benefits on both computing and financial levels. The authors in [22], for example, emphasized the importance of integrating cloud computing in smart cities and suggested a model for this target (see figure 3). The model consists of several hierarchical layers that are responsible for the mapping between the smart cities' requirements and the offered cloud services.

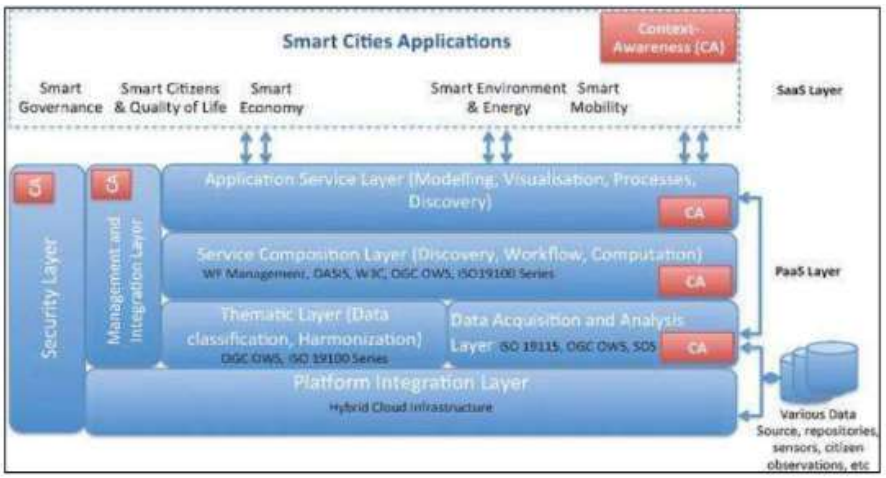

Figure 3: Cloud Computing in Smart Cities [22]

The layers in the model could also be divided into horizontal and vertical sections. The horizontal layer will be responsible for establishing the interfaces between the different components of the smart city and the cloud infrastructure. On the other hand, the vertical layer will provide the necessary security and data management actions before committing the data transactions on the cloud. This model facilitated the process of uploading the large data files to the cloud servers and retrieving them when needed. In addition, it offered more availability and easier interaction for the users in accessing different city modules.

In [23], the authors provided a dense study about the use of cloud services as a revolutionary step in the healthcare sector. The paper included studying the strengths and weaknesses of the cloud when dealing with real-life healthcare applications. Figure 4 illustrates the main strengths, weaknesses, and threats included in the research.

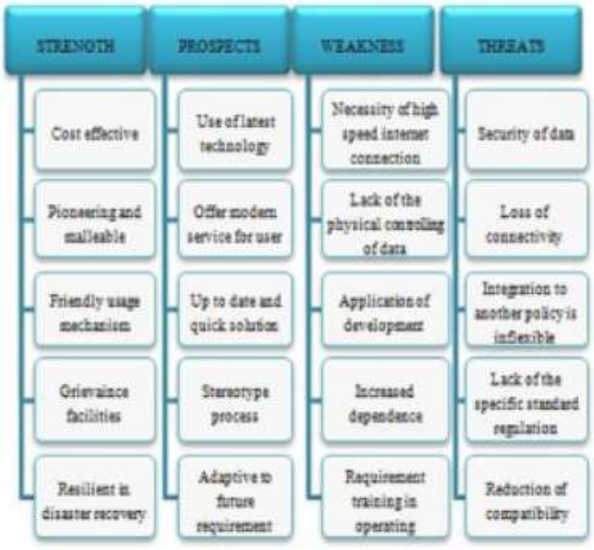

Figure 4:Strengths, Weaknesses, and Threats of Cloud Computing [23]

In addition, the paper tackled some of the cloud models and their properties. For instance, the public cloud is usually accessible by all the users and owned by large entities like governments. On the other hand, the private cloud belongs to a specific institution that provides a set of computing services and management.

Another paper [24] investigated the role of the cloud in agriculture. It highlighted the main weaknesses in the traditional information system that could slow the production chain and limits the tracking process. Those weaknesses include the old unmodernized programs that cannot satisfy the modern agricultural demands, which require enormous data storage and effective processing power. In addition, this lack of a competent information system leads to losses at the financial level and wastes the farmers' efforts.

Then, the authors introduced a cloud architecture (see figure 5) that facilitated the production process and improved the flexibility of the system. For example, such a model provided a better weather tracking system which is critical for the production process.

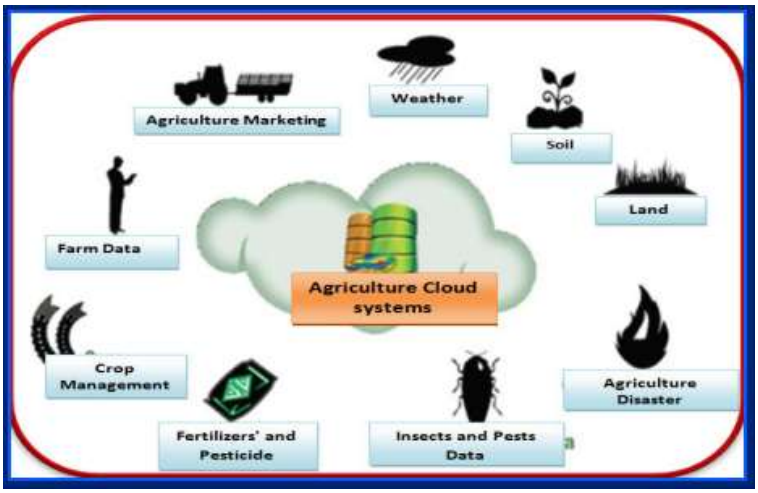

Figure 5: Cloud Computing in Agriculture [24]

The proliferation of IoT devices and real-time applications, which require a lot of computer resources (i.e., storage and processing power), led to the invention of fog computing as a way to enhance cloud services. The research paper [25] was one of the first papers to tackle fog computing and its important role in IoT device development (see figure 6).

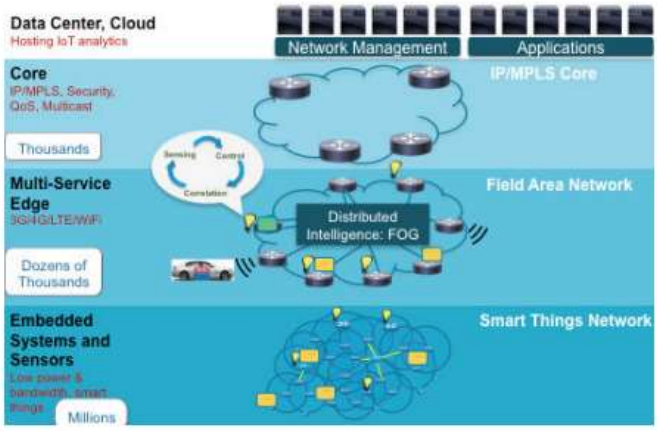

Figure 6: Fog with IoT Structure [25]

The work presented the characteristics that allowed fog to play this role, which includes layer location, decentralization, number of nodes, and real-time communication.

- Layer location: The fog layer is in the middle between the cloud and IoT devices, which helps in filtering data to decrease the pressure on the cloud.
- Decentralization: The decentralized architecture will allow serving the wide distribution of IoT devices.
- The number of nodes: A large number of nodes will eliminate the single point of failure problem that exists in cloud services.
- Real-time communication: The fog location, which is near the end devices (IoT and users), will improve the response time, especially in applications that require real-time data exchange.

The researchers in [26] focused on fog computing as a middle layer and suitable architecture for real-time application. The location of the fog layer helped in being a filter for the data uploaded to the cloud. Hence, instead of uploading data directly, fog nodes can remove the unnecessary information, thus decreasing the pressure on the cloud.

In addition, the authors suggested a model that provides a real-time response based on fog computing (see figure 7). The model presented a unit called "Third Party Memory Management", which is responsible for serving the IoT devices in real-time without the need to use the cloud. In addition, it implemented a mechanism that differentiates between regular and real-time requests.

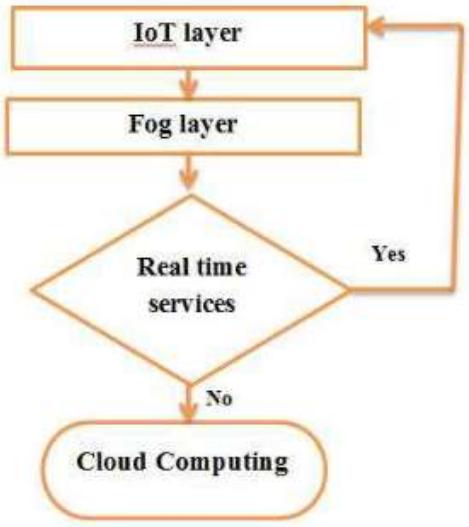

Figure 7: Real-Time Fog Model [26]

In [27], the authors highlighted the importance of the fog layer in the application of the autonomous car because of the need to get a fast response time and less latency. This model involves the cooperation between multiple fog nodes at the same time to serve a set of cars (see figure 8).

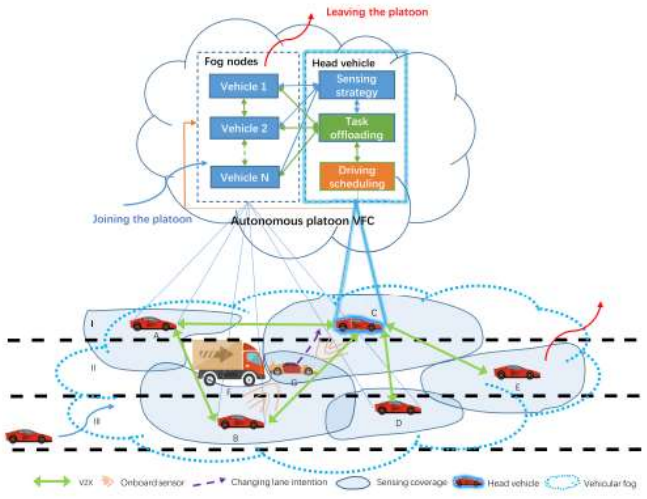

Figure 8: Fog Computing with Autonomous Cars [27]

The model integrated the Support Vector Machine algorithm with fog computing to provide fast response and better cooperation between the different cars. In addition, the decentralized structure of the model helped in selecting the trajectory of the cars on the streets. Finally, simulation results on a real-life dataset showed promising results.

Fog Computing in healthcare Systems:

Many studies have adopted fog computing as a model for healthcare systems. The authors in [28] suggested a fog model that manages a real-time notifications system about the patient's health record. Such a system had low latency and fast response time to avoid querying outdated information. They mentioned that cloud computing cannot handle such an application because of the overhead produced by enormous data. The authors adopted the model in figure 9, which is divided into four layers. The sensor layer consists of the IoT devices and the sensors that gather data from the environment. The fog layer will be responsible for the data analysis and other operations that require a fast response time. The cloud layer will store the large data files. Finally, the manager's layer will monitor and benefit from the model.

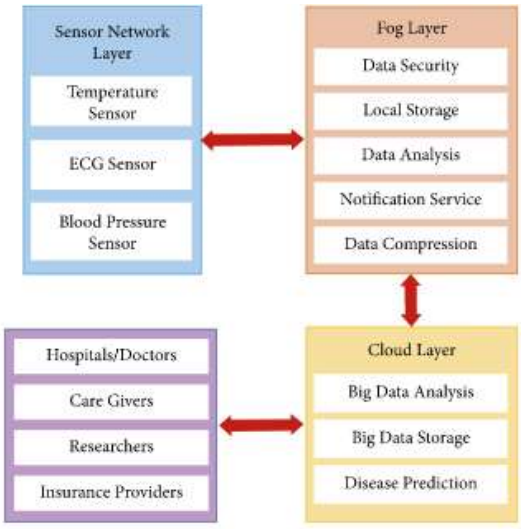

Figure 9: Healthcare Monitoring Architecture

In [29], there was a comparison between a cloud model and a suggested fog architecture in the healthcare system based on security and performance levels.

The model, in terms of services, was divided into two functionalities. The first one, which requires a fast response time, is directly connected to the fog layer (see figure 10) to ensure fast data manipulation. On the other hand, the services that require large storage and high processing power; are directly connected to the cloud.

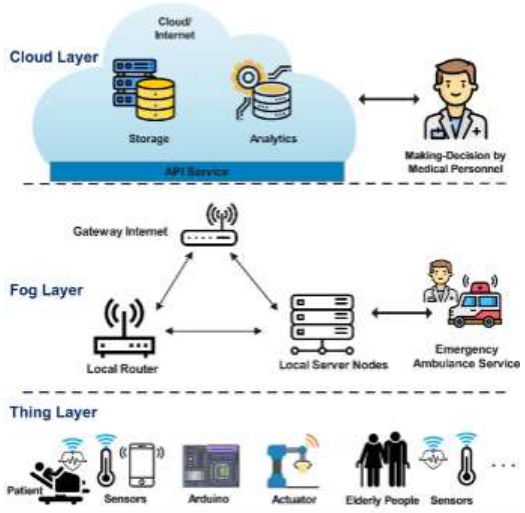

Figure 10: Fog Architecture in Healthcare Systems presented by [29]

On the performance level, the model showed its superiority with 28%, faster than the cloud model. While on the security level, the decentralized structure brought an effective defense against some attacks.

In the paper [30], the model enhanced the presented fog structure by adding a security mechanism that can protect the fog nodes by authenticating the patients who are using the system (see figure 11). In addition, they considered Virtual Machine (VM) in the selection process of fog nodes for better IoT management. Moreover, they suggested a cryptographic model that monitors and supervises the generation of the public and private keys in the system.

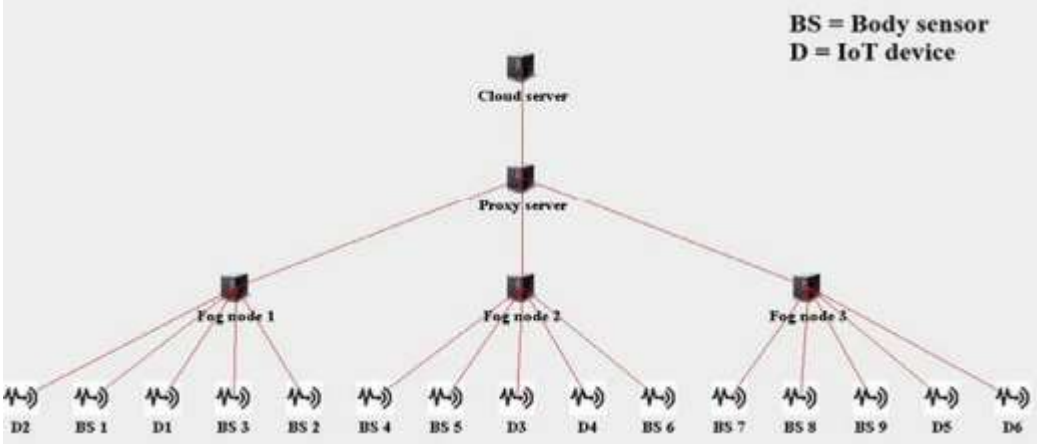

Figure 11: Fog structure in [30] to Ensure Secure Communication

The experimental results were simulated on iFogSim, which is a software for simulating the fog computing structure. Results showed good latency value and system performance.

2.2 Fog Computing Models with Security Solutions

As stated earlier, securing the fog systems could be divided into two categories "Detection and Recovery" and "Prevention". Researchers extensively studied both cases and suggested a lot of tools and models. In this section, we will go over the related work, especially the ones focusing on blockchain solutions in fog environments.

2.2.1 Detection and Recovery

In [2], the authors studied the data integrity in fog systems in healthcare systems. This research focused on the malicious transactions that could spread throughout the system, causing a lot of damage. For example, they stated that a small piece of corrupted data could lead to inconsistent data at multiple fog nodes because of the interrelated database queries. Their study took into consideration both homogeneous and heterogeneous fog architectures (see Figures 12 and 13). The homogeneous structure has a moderator node that can access any

data in the system; thus, helping the doctors to view and edit their patients' records when needed. On the other hand, the heterogeneous structure allows the exchange of information directly between the different nodes without a master, exposing fast inconsistent data spreading.

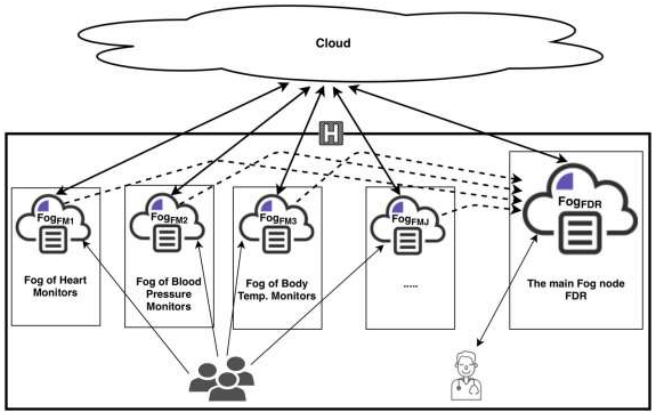

Figure 12:Homogeneous Healthcare Architecture [2]

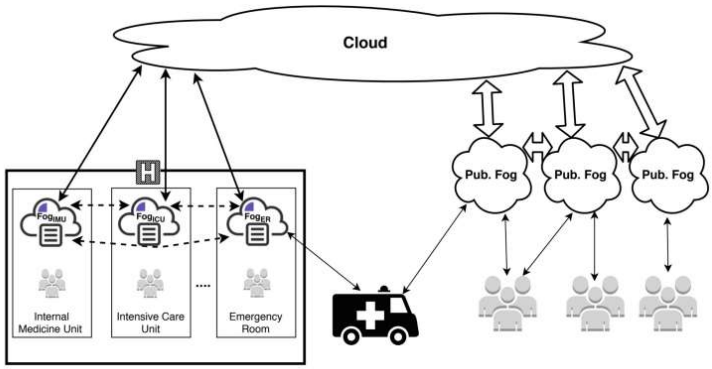

Figure 13: Heterogenous Healthcare System Architecture [2]

The proposed assessment algorithm relied on an intrusion detection system (IDS) to detect the set of malicious transactions in both structures. Next, it searches the possible places of spreading. For example, the affected data could be in the cloud, or it is still in one of the fog

nodes. Then by a procedure that involves searching the history of transactions of the fog nodes, the system will finally assess the amount of spreading. After finishing the assessment process, which will locate the places of affection, a recovery step will be done.

The importance of the presented algorithm is in the capability of assessing the damage in the database efficiently. Moreover, unlike the previous work, after the assessment process, the affected transactions will be distinguished as bad transactions instead of deleting them. This will help in benefiting from them for future work.

Although the paper presented a clear assessment model, it lacked the real-world simulation, which is capable of demonstrating its ability to get a scalable solution. Also, the system relied on the result of the IDS, which may be inaccurate, leading to a false database assessment.

Another study [5] suggested a new fog model for managing users' data in smart cities (see Figure 14). The model presented two algorithms for data assessment and recovery in case of attacks. The model differentiated between two types of fog nodes (private and public) to serve the different computing duties in the city. The private nodes, which are not accessible by the citizens, will manage the public utilities in the city like water, electricity, etc. On the other hand, the public nodes will serve the people and provide the link to the private nodes.

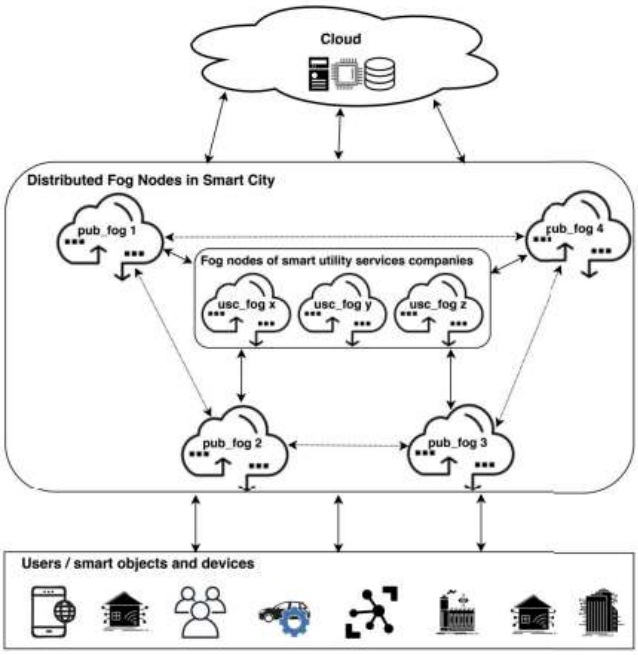

Figure 14: Fog Computing Model in Smart Cities [5]

Then, the authors developed an algorithm to evaluate the damage to the database after being attacked. This algorithm will scan the different nodes in the system and locate the affected transactions relying on the generated logs. Finally, a recovery algorithm will fix the database by redoing the affected transactions. No simulation was done at the end to test the model.

Starting from the fact that authentication alone can not protect the fog computing environment because of malicious transactions from authorized users, [31] proposed a model to isolate the affected fog nodes and reduce the damaging effect. The new model, which is named COMITMENT, imposes a set of procedures to ensure a specific quality of service and a safe environment to exchange data. It is equipped on every fog node and can monitor several services at the same time. For security reasons, the software's algorithm builds trust records and labels the trusted and malicious fog nodes. On the performance side, COMITMENT manages the amount of data processed at each node; thus, enhancing response time. The

simulations showed that the model decreased the intensity of the attacks by 66% and reduced the average latency by 15 sec.

The authors in [32] targeted the detection of malicious nodes as a way to mitigate the damage. The mechanism to detect those nodes relied on an algorithm that studied the behavior and correlation between the different fog servers. For example, an unusual action or communication between two fog nodes could be a signal that there is an attack. In addition, it uses a rating method in which reaching a specific threshold raises a warning. This model proved its importance and effectiveness in the field of vehicles ad hock networks in fog topology.

Similar to [32] in the method to detect malicious nodes, the authors in [33] presented the model "DataIDS" that is responsible for detecting the attacks on the database of the fog nodes. Such a model showed a data analysis methodology to generate dependency graphs and catch any unexpected behavior in the system, which can be evidence of an attack. The simulation of the suggested model showed an effective response toward the noise, replay, and stuck-at attacks. Although the paper showed good outcomes in the experimental results, it didn't go over any recovery approach to cancel the effects of the attack.

A survey presented in [34] studied the integration between machine learning and artificial intelligence from one side with the security of the fog computing system that uses IoT devices. The survey showed that machine learning algorithms were widely used as an effective technique to secure the data in fog computing and protect it against different types of attacks. For example, the Random Forest algorithm assisted in the assessment of the damage after malicious actions. Naive Bayes provided probability models to detect the attacks on the

database. Principal Component Analysis (PCA) algorithm was widely adopted in building intrusion detection systems to catch unusual actions.

2.2.2 Prevention

Prevention solutions in fog computing mainly focus on blockchain technology because of the distributed ledger structure that stops the attacker from committing any transaction without being approved by the majority of the chain's users. In addition, blockchain provides a private environment where different entities in the chain can exchange information without being tracked. Thus, hackers will find an additional problem in understanding the communication architecture of the system. This sub-section will provide the latest and most important studies that tackled this situation.

The researchers in [35] presented a blockchain architecture for fog computing to manage the healthcare records generated by the IoT devices. Their study focused on inserting a new fog layer to increase the capabilities and performance of the suggested model. The system eliminates the latency and throughput problems caused by the direct use of IoT devices. The blockchain ledger was built over the fog layer, forming what they called FogChain.

The model supports real-time data exchange and provides a suitable response time for many types of applications (see figure 15).

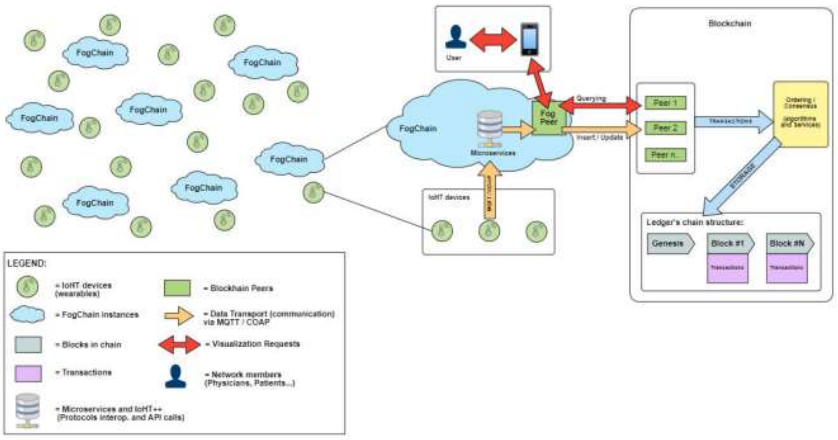

Figure 15: FOGCHAIN architecture presented in [35]

Results showed that the suggested model was better than the cloud model in the number of performed transactions and response time. The model achieved 66% better response time in comparison to the default cloud architecture.

The authors in [19] tackled the case of integrating the fog computing and IoT with the evolving blockchain technology. The author suggested a model that uses fog nodes as users in the blockchain ledger to provide a secure environment for communication and exchanging data. Moreover, they presented an algorithm for IoT devices to help in the process of sharing information and creating new blocks in the chain. The algorithm's main operations were searching, selecting, and examining the environment. Moreover, they were the main metrics to test the performance of the model. In the end, the results showed good performance. Although the paper had promising results, it did not take into consideration the scalability factor that could eliminate the advantages of the system. In addition, the author evaluated the system performance using the local parameters, which may be insufficient to get an accurate result.

In [21], a blockchain-based fog model was proposed to provide a secure environment for data exchange. The paper claimed that the main source of the threat was the data produced from

the IoT devices at the patient level. For that reason, they suggested a structure that uses blockchain technology and encrypted digital signatures to protect the data at the fog nodes from being illegally modified. Moreover, the transparency of the blockchain transactions enhanced the fog computing defense and made tracking modifications easier. The experimental simulations relied on two metrics response time and space complexity. Results showed that the model had good performance on security and scalability levels.

In the last few years, smart cities models have become a field of interest for many researchers. The authors of [20] developed a new structure for smart cities based on fog computing and blockchain technology to enhance the security features and to get a better system performance. The model consisted of a set of layers (Fog, IoT, and Cloud) that cooperate and exchange information to satisfy the different needs of the users. The blockchain, which was on top of the IoT layer, provided the necessary protection against malicious transactions. In addition, the model adopted a set of encryption and authentication steps to ensure the identity of the users before accessing the system. Experimental results, which used the IFogSim software, showed promising results in terms of security, response time, and energy consumption. Despite the advantages of the model, the authors did not go over the scalability factor, which could greatly affect the performance.

With the increase in the number of users in the blockchain ledger, the technology suffered from a decrease in performance, causing a lot of delays in creating new blocks. For that reason, many studies tried to achieve scalable and secure chains that can meet the user's requirements. In [8], the researchers integrated the fog computing environment with an improved version of the blockchain, which is sidechains. Sidechains partitioned the whole system into several chains

that can work independently but with the cooperation of a root chain. It is worth mentioning here that those sub-chains may share the same structure and architecture or may not. In the simulation, the authors used the plasma software, considering some low-power IoT devices as users in the chains. In addition, they used Rasberry pi hardware to play the role of the clients. Experimental results showed that the new model was able to achieve better utilization of the processing power and get more successful transactions.

Another paper [17] focused on the sidechain solution as a way to overcome the scalability problem in fog computing. This study presented an architecture that is capable of combining the work of several sidechains with the coordination with the root chain. In addition, it supported an access mechanism to regulate the work of the different ledgers and approve the set of generated transactions in the system. The simulation was done through IOTA Tangle software with many metrics such as throughput, latency, and sending rate. Experimental results showed the scalability and efficiency of the model compared to the regular blockchain.

Chapter 3

The Suggested Model

This section of the thesis will present the model and the approach adopted in this work. As previously mentioned, fog computing architecture showed promising performance in applications that require strict quality of service constraints. In the field of healthcare, those constraints are the latency and the response time because of the importance of getting updated information about the patient's records. Moreover, any outdated or misleading data could put lives at risk. In addition, some healthcare applications require a consistent notification system that sends accurate and precise information.

Although fog computing presents an effective and efficient environment for data exchange in the healthcare system, it is still vulnerable to malicious transactions that could spread out through the whole system, causing a lot of corruption. The suggested model will solve this problem by using blockchain technology that is capable of rejecting untrusted transactions based on the approval of the available fog nodes. Moreover, the model will overcome the scalability problem with its clustering and sidechain technique.

3.1Model/Overview

The model will start first by clustering the fog nodes based on the frequency of communication and with a number k that specifies the minimum number of nodes in a cluster. This step will help in locating the fog nodes that regularly exchange data between each other. Thus, providing

local privacy between those nodes. The k-means clustering algorithm [36] is adopted in the model, and a number N will specify the number of clusters. Before running the clustering algorithm, the least 10 interacting fog nodes will be excluded and will form the main chain of the system.

We chose number ten to form the main ledger because it is widely recommended that the minimum number of nodes to form a secure blockchain is seven. The recommendation was based on the fact that seven nodes provide more than 66% agreement between nodes with the ability to tolerate two untrusted participants.

The k-means clustering algorithm will work based on the number of communications between the different nodes and will cluster them accordingly. It is clear that the more links between two nodes, the more likely they will be in the same cluster when the algorithm converges. Then the formed clusters will form the sidechains, and each node within the cluster will represent a user in the chain. The fog node in a specific cluster will be responsible for committing the transactions for all the connected IoT devices. In addition, the transactions within each sidechain will be monitored and approved by the nodes of the cluster only. Such a step can provide more privacy for the entities that require a level of secrecy and doesn't want their data to be visualized or tracked. The creation of the sidechains is done through the plasma framework, which initializes a ledger with a given number of transactions and specific data to be processed. Moreover, all the manipulation process (adding blocks) is based on the Ethereum protocols that involve specific features like smart contracts, Decentralized Autonomous Organizations (DAO), and digital token. Figure 16 illustrates the mentioned steps.

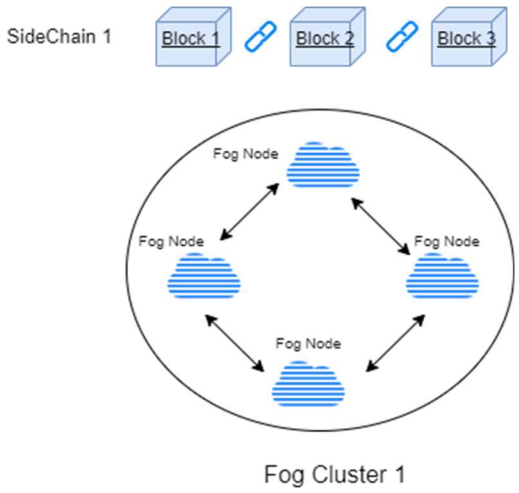

Figure 16: A Fog Cluster

Each block consists of a set of parameters that are essential for its functioning (see figure 17):

Block ID (same as Transaction ID): is a unique attribute that refers to a specific block.

Hash: As stated in chapter 1, each block has a unique hash that ensures that the block is not altered or modified by unauthorized users.

Previous Block Hash: This parameter creates the ledger structure by linking the different blocks to each other. When a block is modified, it will change its hash, thus making the link inconsistent. For that reason, it is considered the main defense mechanism in the blockchain.

Encrypted content: The healthcare data is encrypted and saved in this attribute of the block.

Signature: This attribute links the creation of the block to a specific entity without knowing the actual identity.

Timestamp: Records the date of the creation of the block.

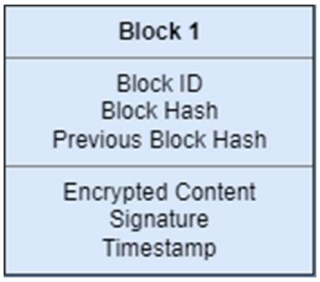

Figure 17:Block Structure

The main chain will consist of the fog nodes in the system that do not belong to any cluster or preselected ones. This chain will be responsible for exchanging the transactions between the different sidechains. The communication between the sidechains and main chain will be through a protocol called a 2-way peg, which ensures the integrity of the data transmitted between them. This protocol is the most important component in cross-chain data exchange because it ensures the correct communication and information transfer from one chain to the other. Moreover, it obliges both chains through a digital contract to abide by the confirmed data transactions. Similar to side chain creation, the main chain involves the same set of features provided by the plasma framework. The main chain is built based on the data provided for the ten selected nodes and allows the cross-chain data exchange using the 2-way-peg protocol.

Coordinator

The coordinator presented in the model is a computer program that is responsible for the encryption/decryption process to ensure that the data is only accessible by authorized users. Moreover, it plays a role in helping the recipient node to find the intended data after being uploaded to the main chain.

Encryption/Decryption process

This process is done based on the private and public keys concepts that can protect the data from unauthorized access. The public key will be shared between all the nodes and has the role of encrypting the data. On the other hand, the private key is given to specific nodes that have the right to decrypt the data (see figure 18).

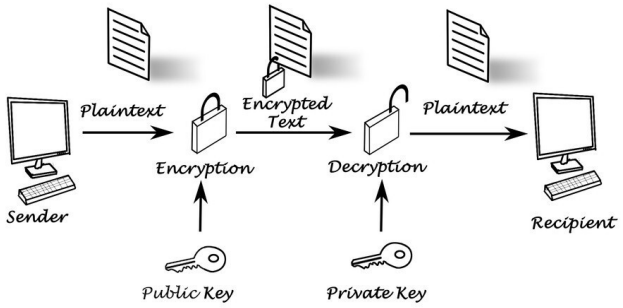

Figure 18: Public and Private Key Concept

First, the node uploading the data will encrypt it using the public key provided to all the nodes (sent by the coordinator). In the next step, the coordinator will follow a predefined set of privileges to send the private key to the authorized recipients to ensure their right in accessing the data.

Implementing the different chains in the model is done through the plasma framework [37] that allows the available active nodes to be divided into chains. The framework creates the chains using an interface that permits adding the nodes to each chain (including the main chain) and setting the different attributes like proof of work, time to approve a block, and the data exchanged. Moreover, this software specifies the data exchange scheme that will select the data used in the model and the flow of data and transactions between the different entities in

the simulation. The importance of this framework is in the 2-way-peg protocol, which ensures a smooth data transfer between any two chains. This protocol works by locking the data on the sending chain first and then using a smart contract, the data will be transferred without any modification to its intended destination.

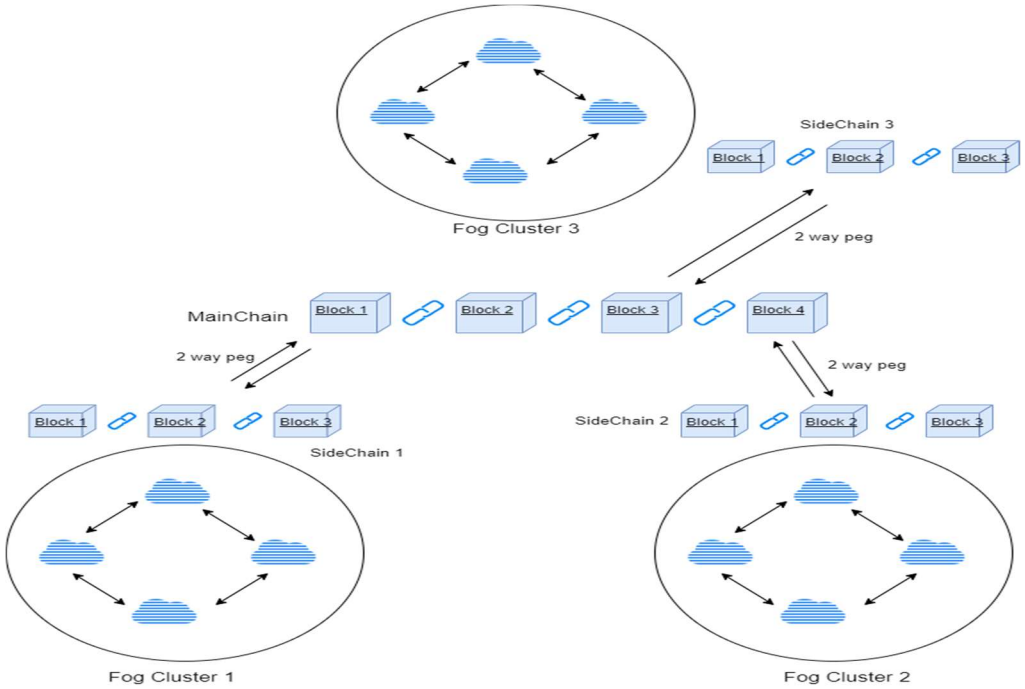

Figure 19: The Complete Model

Pseudo Code

This subsection will present the pseudo-code for the creation of the whole model. It will include the clustering method, the creation of the chains, and the data mapping to the chains.

1. k // minimum number of nodes per cluster
2. N // number of clusters
3. D //data
4. Exclude the 10 least active nodes from D
5. Run k-means (D, k, N)
6. Establish SideChains + MainChain
7. Map the data to the different chains;
8. Run the transactions

3.2 How the model works?

3.2.1 In the sidechain of a specific cluster

In a given cluster, when a node commits a transaction, a block is created containing the encrypted data with a specific hash and points to the hash of the previous block. This process is done using the plasma framework that links the different blocks and updates the ledger at each node within the cluster. This block is then sent to the other fog nodes in the same cluster to check if it is accepted or not (All/Majority of the users). Finally, if the approval is achieved, the block is added to the sidechain of all the users (in the same cluster); else, it will be deleted.

Pseudo Code

The below pseudo code will present the functionality of the sidechain in a specific cluster, including the approval mechanism and the nodes creation process.

```
1.  Di → database of Fog node i
2.  Tj → Transaction j
3.  Li → SideChain ledger of node i
4.  Bj → Block of transaction j
5.  If Tj.Is_committed(Di)
6.          Bj.Create();
7.  If majority of nodes approve
8.          Bj.Add_To(Li);
9.  Else
10.         Bj.Remove();
```

3.2.2 Exchanging data between clusters:

When two fog nodes in two different clusters need to exchange information, the data will be sent first to the main chain through the 2-way-peg protocol. This protocol is responsible for ensuring the integrity of the data while being transferred from one chain to the other.

Moreover, it will lock the data in the side chain and wait until the smart contract is initiated to transfer it from one chain to the other. Then, from the main chain to the targeted cluster (also using the 2-way-peg), the fog node could get the data. It is worth mentioning here that at every included chain, a new block will be added to the ledger.

Pseudo Code

This part shows the functionality of the whole model when communication between the different clusters is involved. It will include how the data is transferred from one block to the other and blocks creation locations.

```
1.  Di → database of Fog node i
2.  Dy → database of Fog node y
3.  Tj → Transaction j //sending data to main chain
4.  Tk → Transaction k //receiving data from the main chain
5.  Tm → Transaction m // posting data on the main chain
6.  Lm → Main Ledger
7.  Li → SideChain ledger of node i
8.  Ly → SideChain ledger of node y
9.  Bj → Block of transaction j
10. Bk → Block of transaction k
11. Bm → Block of transaction m
12. If Tj.is_Commited(Di)
13.       Bj.Create();
14. If majority of nodes approves (Cluster i)
15.        Bj.Add_To(Li);
16. Else
17.        Bj.Remove();
18.        End_Process();
19. If  Tm.is_Commited(Di)
20.       Bm.Create();
21. If majority of nodes approves (Main Chain)
22.       Bm.Add_To(Lm);
23. Else
24.       Bm_Remove();
```

```
25.     End_Process();
26. If Tk.is_Commited(Di)
27.     Bk.Create();
28. If majority of nodes approves (Cluster y)
29.     Bk.Add_To(Ly);
30. Else
31.     Bk.Remove();
32.     End_Process();
```

3.3 An Example

In this sub-section, we will demonstrate how the suggested model works in the case of intercluster and intracluster communication.

3.2.1 Data exchange within a cluster:

After taking some medical measures for a patient called “Jad”, the responsible department decided to upload the data into the sidechain to be visible to the others. The table below shows the records of the patients.

ID	Name	Temp	Weight
1234	Jad	38	70

Table 1Records number 1 of Jad

As a first step, the fog nodes responsible for the operation will create a new block containing the encrypted data of Jad and a hash corresponding to it. Then, if the majority of the nodes approve the block creation, it will be added to the ledger of the cluster. The votes are managed and organized using the plasma framework that is capable of gathering the votes and giving the decision of the new block.

Let's suppose that the transaction of Jad was successful. Now a doctor from another fog node within the same cluster needs to access the data. He needs to request permission for the data (from the entity which uploaded the data) through the coordinator, who will provide the necessary private decryption key. After obtaining the private key, the node will be able to decrypt and use the content of the intended node on the distributed ledger of the sidechain.

3.2.2 Data exchange between clusters:

In a hospital, the emergency department is in fog cluster 1, and the X-Ray entity is in fog cluster 2. A patient called Sami entered the emergency and had the following record:

ID	Name	Temp	Weight
4321	Jad	37	75

Table 2Records number 2 of Jad

After transferring the patient to the X-Ray department, doctors in that department requested the old data.

Fog cluster 1 will request creating a new block on the main chain containing the data (with the help of the coordinator). Then, the coordinator will be responsible for exchanging the crypto-keys between the targeted fog nodes. Using the 2-way-peg protocol, the sidechain of the first cluster will add the data to the main chain through another cross-chain operation. The second cluster will be able to get the data to its ledger. The data is finally decrypted using the private key provided by the coordinator.

Chapter 4

Experimental Results

This part of the thesis will examine the experimental results after the simulation of the suggested model. It will first introduce the simulation software which was used to get the results. Then it will go over the dataset. Finally, it will show the achieved results.

4.1 Simulation Software

The Plasma framework was used to simulate the functionality of the suggested model. This framework supports InterLedger Protocols, which play a vital role in the data exchange between the different chains. Truffle and Ganache were also used to allow the smart contract in the environment.

4.2 Data

We used simulated data for fog computing in the healthcare systems to test the suggested model. The dataset consists of medical reports for clients, assuming that the records may be X-ray photos, medicine prescriptions, or any type of text [7]. In the model, we will assume that the block in the sidechain and mainchain can have a single medical report only. The maximum size that a block can handle is 2.53 KB.

4.3 Hardware

In the simulation, a 2.30GHz computer with 8 GB of RAM, a core i7 processor, 1 TB hard disk, and a 64-bit windows operating system was used to run the Plasma framework.

4.4 Performance

This subsection will focus on presenting the results after doing the simulation on the Plasma framework. The metrics “Transaction per minute”, “Latency”, and “Response time” will be used to study the performance and scalability of the model in comparison with the default blockchain approach. Moreover, the data exchange size will be tracked as well. In addition, a discussion of the achieved results will be discussed and analyzed.

In figure 19, the graph shows the performance of a single chain in comparison with five sidechains. It focused on studying the number of committed transactions as a function of time. Results clearly show that the model that contains the five chains is superior in terms of productivity, which lead to more transactions in the given time. The clustering technique, which decreased the number of nodes in each one of the sidechains, made the data exchange smoother and faster because the approval of the transactions is required by a smaller number of nodes. In addition, the proof of work, which is an elementary protocol in providing the securing and validity in the chains, requires less running time in comparison to having a single chain model.

In figure 19, at minute two, we can spot the little curve in the five chains model; this is justified by the fact that the simulation is on a single computer that may encounter a varying performance depending on the RAM and CPU utilization.

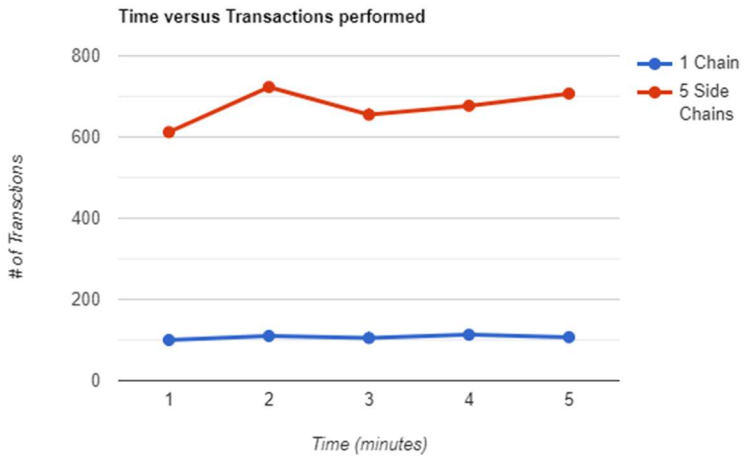

Figure 19: Time versus Transactions performed

Figure 20, on the other hand, studied another Latency which is an important metric when assessing the effectiveness of a blockchain mode. This metric is defined as the time from the submission of the transaction until it was successfully rejected or accepted on the chain. The graph in the figure clearly showed less latency for the new model because of the fewer number of transactions per chain, so the transaction time to be approved or disapproved will be less.

Although the size of the block is 2.53 KB, in the 5 clusters model, we can spot that the latency decreased because of the less pressure on the chains, which leads to more data exchange and less latency for the transactions.

Similar to the previous figure, we can spot the small change in performance and a curve at minutes two and four because of the change in the computer processing power.

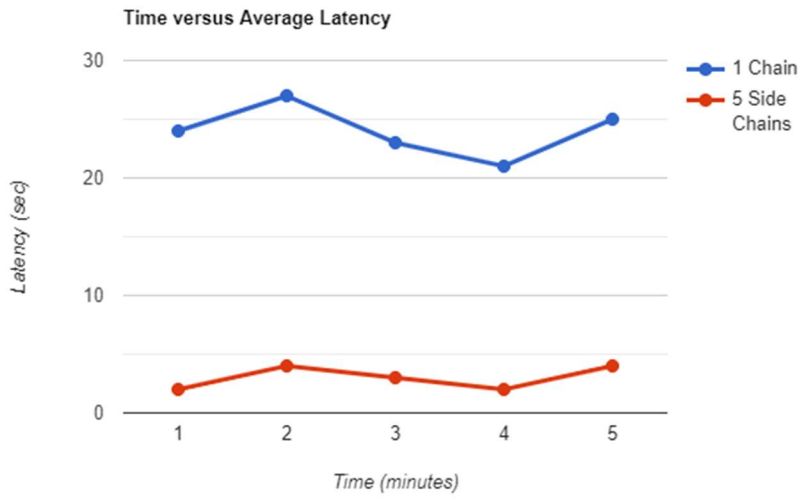

Figure 20: Time versus Average Latency

In the simulation, this thesis studied the size of data exchange through the model. The data exchange is the amount of data that went through the blockchain using the transactions. Figure 21 shows the data that entered the two models as a function of time. It is clear that the new model allowed much data to enter the chains. This superiority could be explained by the less latency and more transactions that are being approved to the suggested architecture. The clustering technique allowed more data flow throughout the model because it eliminated the single chain structure, which proved to allow less data exchange when a large number of entities are manipulating data.

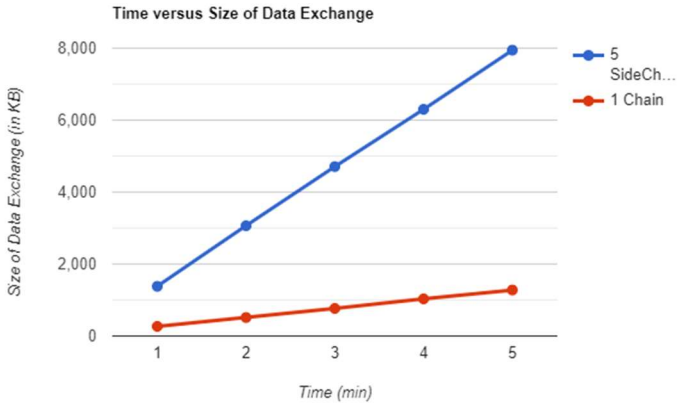

Figure 21: Time versus Size of Data Exchange

After studying the performance of the model, figures 22 and 23 tackle the scalability when increasing the number of clusters to a large number. The model was tested when the number of clusters was 30 to check its ability and effectivity in adding blocks to the ledgers. It is clear that the increase in the number of clusters positively enhanced the performance of the system in terms of the data exchange and the number of performed transactions. This promising result proves the importance of the model as a scalable solution for fog computing because of the ability to deal with information systems with a high number of nodes.

The clustering technique eliminates the single chain structure that highly affects scalability and requires the validation process for a huge number of nodes. Moreover, it takes the advantage of the sidechain architecture that requires less effort in block creation, which leads to better system throughput. In addition, the sidechain does not require permanent connectivity with the mainchain, which avoids the bottleneck problem.

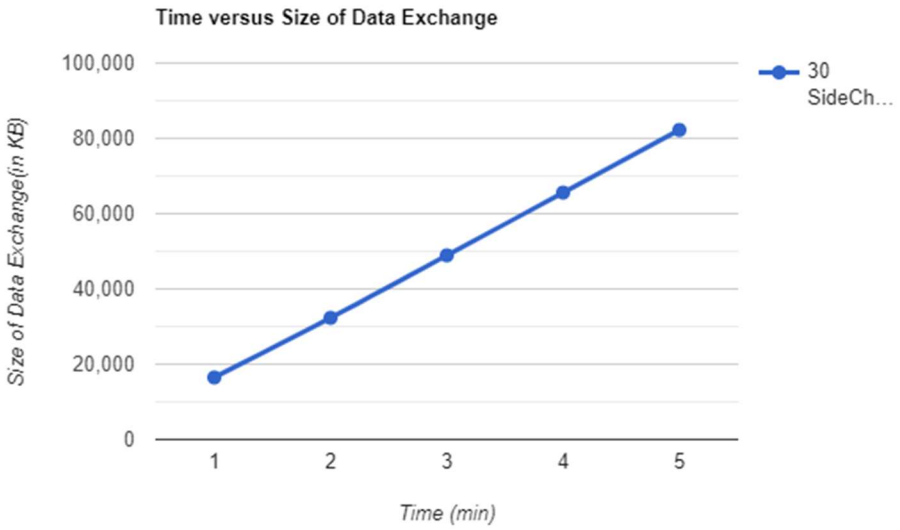

Figure 22: Time versus Size of Data Exchange

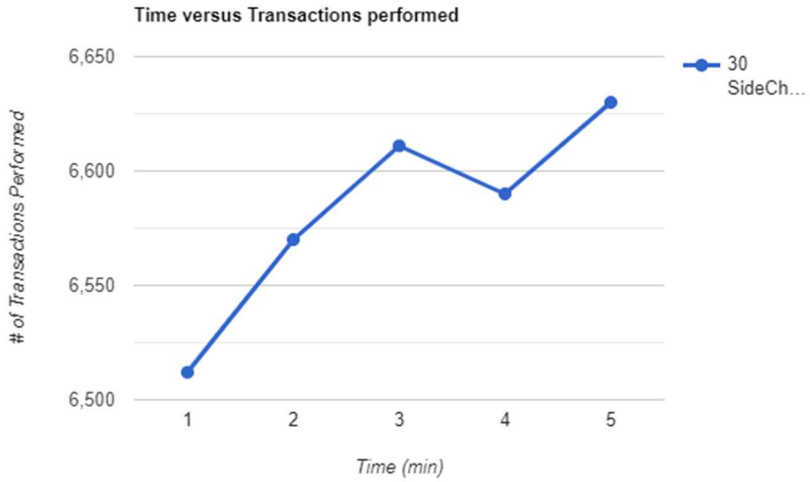

Figure 23: Time versus Transactions performed

We compared our model to the system presented in [35]. In this study, the authors showed a novel architecture in securing healthcare records in a fog environment. The model focused on a new fog layer to secure the system and provide better throughput and real-time services.

Because the model in [35] adopted a block of the size of 1KB and 0.1KB, we performed the simulation on the same block sizes. Figures 24 shows a comparison between our approach (30 clusters) and the system presented in [35] in terms of throughput (Transactions per second). The bar graph clearly shows that our model is superior in terms of the number of transactions performed; This could be justified by the clustering algorithm that forms the side chains and decreases the validation time required by each transaction.

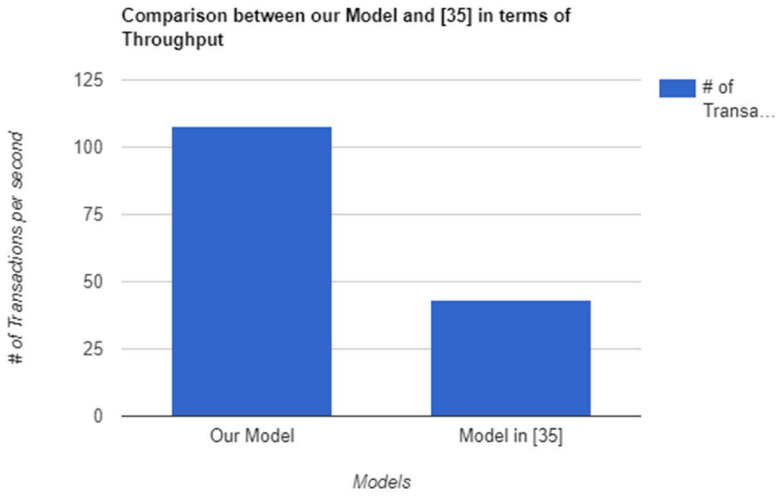

Figure 24 Comparison between our Model and [35] in terms of Throughput

In addition to throughput, we could compare our model with the latency results presented in [35] and the latency of the cloud model. Results presented in figure 25 shows the superiority of our model in term of latency. The sidechains with their smaller node numbers can allow the transactions to be committed faster and decrease the time required for the validation process.

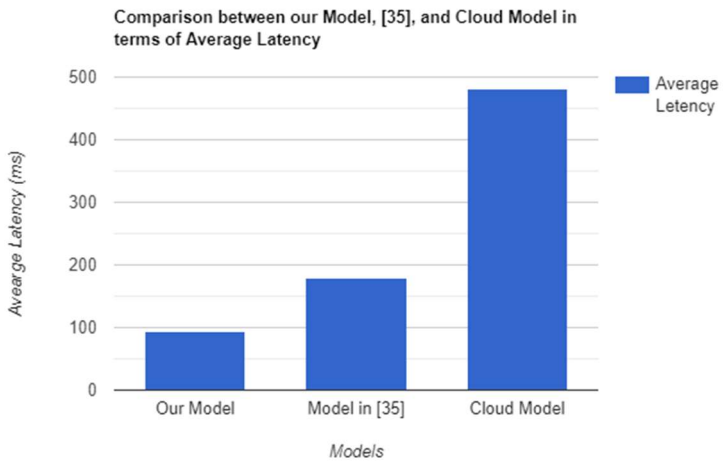

Figure 25: Comparison between our Model, [35], and Cloud in terms of Latency

Chapter 5

Conclusion and Future Work

Cloud computing, with its services enhanced the performance of information systems. Those services allowed the users to benefit from better processing power, more options, and enormous storage. However, with the prosperity of IoT devices and the increase in complexity and demands of the programs, the cloud was not able to satisfy the user's needs. For that reason, the fog layer came as an enhancement to the cloud and offered the users a variety of advantages, including fast response time and less latency. Despite the superiority of the fog, it is still a target for attackers who commit malicious transactions between the different nodes.

Based on that, this work suggested a fog model based on a clustering algorithm to form a set of sidechains for monitoring the flow of transactions throughout the system. In addition, it addressed the scalability problem, which is one of the common drawbacks of blockchain technology and could eliminate its advantages.

The security in the model comes from the hash structure that protects the nodes from being altered. This defense mechanism forces the attacker to change the hash of all the preceding blocks at all the users, which seems impossible. In addition, the decision mechanism protects the system from injecting bad transactions.

In future work, we could use real data in the simulation process to get more accurate and closer to real-world applications. Testing the model on industrial data helps to detect problems that do not appear when using computer-made data. In addition, this work may

adopt a different clustering algorithm or use another metric than the frequency of communication when clustering. This modification may lead to different discoveries. In addition, for the evaluation process, researchers could use other blockchain metrics in the assessment process. For example, response time could be a metric to study in future research other than transactions per minute and latency that were used to evaluate the model.

References

[1] Elhadad, A., Alanazi, F., Taloba, A. I., & Abozeid, A. (2022). Fog computing service in the Healthcare Monitoring System for managing the real-time notification. Journal of Healthcare Engineering, 2022, 1–11. https://doi.org/10.1155/2022/5337733

[2] Alazeb, A., & Panda, B. (2019). Ensuring Data Integrity in Fog Computing Based Health-Care Systems. Security, Privacy, and Anonymity in Computation, Communication, and Storage, 63–77. https://doi.org/10.1007/978-3-030-24907-6_6

[3] Jiang, Y., Wang, C., Wang, Y., & Gao, L. (2019). A Cross-Chain Solution to Integrating Multiple Blockchains for IoT Data Management. Sensors, 19(9), 2042. https://doi.org/10.3390/s19092042

[4] Yi, S., Qin, Z., & Li, Q. (2015). Security and Privacy Issues of Fog Computing: A Survey. Wireless Algorithms, Systems, and Applications, 685–695. https://doi.org/10.1007/978-3-319-21837-3_67

[5] Alazeb, A., & Panda, B. (2019). Maintaining Data Integrity in Fog Computing Based Critical Infrastructure Systems. 2019 International Conference on Computational Science and Computational Intelligence (CSCI). https://doi.org/10.1109/csci49370.2019.00014

[6] Hosseinpour, Farhoud & Amoli, Payam & Plosila, Juha & Hämäläinen, Timo & Tenhunen, Hannu. (2016). An Intrusion Detection System for Fog Computing and IoT based Logistic Systems using a Smart Data Approach. International Journal of Digital Content Technology and its Applications. 10.

[7] Ismail, L., & Materwala, H. (2020). Blockchain Paradigm for Healthcare: Performance Evaluation. Symmetry, 12(8), 1200. https://doi.org/10.3390/sym12081200

[8] Ziegler, M. H., Grossmann, M., & Krieger, U. R. (2019). Integration of Fog Computing and Blockchain Technology Using the Plasma Framework. 2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC). https://doi.org/10.1109/bloc.2019.8751308

[9] Blockchain Metrics | Arcitura Patterns. (n.d.). ARCITURA. Retrieved July 28, 2022, from http://patterns.arcitura.com/blockchain-patterns/blockchain-metrics

[10] Gia, T. N., Ali, M., Dhaou, I. B., Rahmani, A. M., Westerlund, T., Liljeberg, P., & Tenhunen, H. (2017). IoT-based continuous glucose monitoring system: A feasibility study. Procedia Computer Science, 109, 327–334. https://doi.org/10.1016/j.procs.2017.05.359

[11] Kua, J., Loke, S., Arora, C., Fernando, N., & Ranaweera, C. (2021). Internet of Things in Space: A Review of Opportunities and Challenges from Satellite-Aided Computing to Digitally-Enhanced Space Living. Sensors, 21(23), 8117. https://doi.org/10.3390/s21238117

[12] How Google Workspace uses encryption to protect your data. (n.d.). Google. Retrieved July 21, 2022, from https://services.google.com/fh/files/misc/google-workspace-encryption-wp.pdf

[13] Fatoum, H., Hanna, S., Halamka, J. D., Sicker, D. C., Spangenberg, P., & Hashmi, S. K. (2021). Blockchain Integration With Digital Technology and the Future of Health Care Ecosystems: Systematic Review. Journal of Medical Internet Research, 23(11), e19846. https://doi.org/10.2196/19846

[14] Author, G. (2014, July 31). A Cautionary Tale: How a Bug in Dropbox Permanently Deleted 8,000 of My Photos. PetaPixel. https://petapixel.com/2014/07/31/cautionary-tale-bug-dropbox-permanently-deleted-8000-photos/

[15] Squarepants, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3977007

[16] Foley, S., Karlsen, J. R., & Putniņš, T. J. (2019). Sex, Drugs, and Bitcoin: How Much Illegal Activity Is Financed through Cryptocurrencies? The Review of Financial Studies, 32(5), 1798–1853. https://doi.org/10.1093/rfs/hhz015

[17] Jiang, Y., Wang, C., Wang, Y., & Gao, L. (2019). A Cross-Chain Solution to Integrating Multiple Blockchains for IoT Data Management. Sensors, 19(9), 2042. https://doi.org/10.3390/s19092042

[18] Back, S.A., Corallo, M., Dashjr, L., Friedenbach, M., Maxwell, G., Miller, A.K., Poelstra, A., & Timón, J. (2014). Enabling Blockchain Innovations with Pegged.

[19] Alam, T. (2019). IoT-Fog: A Communication Framework using Blockchain in the Internet of Things. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3638991

[20] Singh, P., Nayyar, A., Kaur, A., & Ghosh, U. (2020). Blockchain and Fog Based Architecture for Internet of Everything in Smart Cities. Future Internet, 12(4), 61. https://doi.org/10.3390/fi12040061

[21] Ngabo, D., Wang, D., Iwendi, C., Anajemba, J. H., Ajao, L. A., & Biamba, C. (2021). Blockchain-Based Security Mechanism for the Medical Data at Fog Computing Architecture of Internet of Things. Electronics, 10(17), 2110. https://doi.org/10.3390/electronics10172110

[22] Agarwal, Neetu & Agarwal, Gaurav. (2017). Role of Cloud Computing in Development of Smart City. International Journal of Science Technology & Engineering

[23] Devadass, L., Sekaran, S. S., & Thinakaran, R. (2017). CLOUD COMPUTING IN HEALTHCARE. International Journal of Students' Research in Technology & Management, 5(1), 25–31. https://doi.org/10.18510/ijsrtm.2017.516

[24] Choudhary, Sushil & Jadoun, R & Mandoria, Hardwari. (2016). Role of Cloud Computing Technology in Agriculture Fields. Computer Engineering and Intelligent Systems. 7. 2222-2863.

[25] Bonomi, F., Milito, R., Zhu, J., & Addepalli, S. (2012). Fog computing and its role in the internet of things. Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing - MCC '12. https://doi.org/10.1145/2342509.2342513

[26] Ema, R. R., Islam, T., & Ahmed, M. H. (2019). Suitability of Using Fog Computing Alongside Cloud Computing. 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT). https://doi.org/10.1109/icccnt45670.2019.8944906

[27] Du, H., Leng, S., Wu, F., Chen, X., & Mao, S. (2020). A New Vehicular Fog Computing Architecture for Cooperative Sensing of Autonomous Driving. IEEE Access, 8, 10997–11006. https://doi.org/10.1109/access.2020.2964029

[28] Elhadad, A., Alanazi, F., Taloba, A. I., & Abozeid, A. (2022). Fog Computing Service in the Healthcare Monitoring System for Managing the Real-Time Notification. Journal of Healthcare Engineering, 2022, 1–11. https://doi.org/10.1155/2022/5337733

[29] Quy, V. K., Hau, N. V., Anh, D. V., & Ngoc, L. A. (2021). Smart healthcare IoT applications based on fog computing: architecture, applications and challenges. Complex & Intelligent Systems. https://doi.org/10.1007/s40747-021-00582-9

[30] Awaisi, K. S., Hussain, S., Ahmed, M., Khan, A. A., & Ahmed, G. (2020). Leveraging IoT and Fog Computing in Healthcare Systems. IEEE Internet of Things Magazine, 3(2), 52–56. https://doi.org/10.1109/iotm.0001.1900096

[31] Al-khafajiy, M., Baker, T., Asim, M., Guo, Z., Ranjan, R., Longo, A., Puthal, D., & Taylor, M. (2020). COMITMENT: A Fog Computing Trust Management Approach. Journal of Parallel and Distributed Computing, 137, 1–16. https://doi.org/10.1016/j.jpdc.2019.10.006

[32] Gu, K., Dong, X., & Jia, W. (2022). Malicious Node Detection Scheme Based on Correlation of Data and Network Topology in Fog Computing-Based VANETs. IEEE Transactions on Cloud Computing, 10(2), 1215–1232. https://doi.org/10.1109/tcc.2020.2985050

[33] Fantacci, R., Nizzi, F., Pecorella, T., Pierucci, L., & Roveri, M. (2019). False Data Detection for Fog and Internet of Things Networks. Sensors, 19(19), 4235. https://doi.org/10.3390/s19194235

[34] Ahanger, T. A., Tariq, U., Ibrahim, A., Ullah, I., Bouteraa, Y., & Gebali, F. (2022). Securing IoT-Empowered Fog Computing Systems: Machine Learning Perspective. Mathematics, 10(8), 1298. https://doi.org/10.3390/math10081298

[35] Mayer, A. H., Rodrigues, V. F., Costa, C. A., Righi, R. da, Roehrs, A., & Antunes, R. S. (2021). Fogchain: A fog computing architecture integrating blockchain and internet of things for personal health records. IEEE Access, 9, 122723–122737. https://doi.org/10.1109/access.2021.3109822

[36] Na, S., Xumin, L., & Yong, G. (2010). Research on k-means Clustering Algorithm: An Improved k-means Clustering Algorithm. 2010 Third International Symposium on Intelligent Information Technology and Security Informatics. https://doi.org/10.1109/iitsi.2010.74

[37] Poon, J., & Buterin, V. (2017). Plasma: Scalable Autonomous Smart Contracts. Scalable Autonomous Smart Contracts. https://plasma.io/