Towards using blockchain technology for IoT data access protection

Nabil Rifi, Elie Rachkidi, Nazim Agoulmine, Nada Chendeb Taher

Towards Using Blockchain Technology for IoT data access protection

[1,2]Nabil Rifi, [1]Elie Rachkidi, [1]Nazim Agoulmine, [2]Nada Chendeb Taher
[1]COSMO, IBISC Laboratory, University of Evry, France
[2]Lebanese University, Faculty of Engineering and Azm Center for Researches, Tripoli, Lebanon

***Abstract*—In the past few years, the number of wireless devices connected to the Internet has increased to a number that could reach billions in the next few years. While cloud computing is being seen as the solution to process this data, security challenges could not be addressed solely with this technology. Security problems will continue to increase with such a model, especially for private and sensitive data such as data personal and medical data collected with more and more sophisticated connected devices (forming the IoT). Thus the need for a fully decentralized peer to peer and secure technology to overcome these problems. The blockchain Technology is a promising approach giving the properties it brings to the field. This paper illustrates an architecture based on blockchain technology, and a protocol for data access, using smart contracts and a publisher-subscriber mechanism.**

I. Introduction

IoT is taking over the world, it is estimated that the number of devices connected to the Internet forming the Internet of Things will reach 50 Billion by 2020 [1]. And one critical application is the eHealth Smart Homes. In fact, this technology allows to monitor elderly or individuals with diseases and automatically send the data to a remote server for processing by doctors. This data is recorded in the so called EMR. Secure access to this EMR is problematic considering security and privacy issues, transparency, etc. This is why in order to develop secure and reliable solutions for eHealth smart homes, it requires unprecedented coordination and collaboration between all pieces of the system. All devices must work together and be integrated with all other devices, and all devices must communicate and interact seamlessly with remote systems and infrastructures in a secure way. Such a solution is possible, however it can be expensive and time consuming. Thus the need of new ideas, new technologies that will converge IoT security towards a decentralized model.
Having this huge amount of data, being centralized, and sometimes monitored by one single provider, may create many issues. The cloud is a computing/storing technology, it cannot protect the security and privacy of its users only by itself. Blockchain technology proved its efficiency in financial applications such as Bitcoin [2] , and it can be of great value and importance. In this paper we discuss first some background and related works. Then we present the specific problem we want to address and the proposed architecture and model. The proposed solution is based on smart contracts [3] and publisher-subscriber mechanism. In the following, we present a mathematical approach and its implementation. Finally, we present a conclusion.

II. Related Works

A. Background

1) Blockchain: In 2008 Satoshi Nakamoto introduced Bitcoin, a fully digital and decentralized cryptocurrency. In order to solve the double spending problem in Bitcoin, Blockchain technology was introduced. It is a peer to peer decentralized distributed ledger that is replicated on all nodes participating in the system. It is a complete transparent technology that can show all the transactions that have been made since its creation, without tampering or fraud. Blockchain is a group of blocks that are connected each block to the one before. The first blockchain block is called the ”Genesis” block and it is hardcoded into the software.

2) Smart Contracts: Smart contracts are codes implemented on the blockchain itself, allowing general purpose computation. Similar to a contract between any two individuals, it can have conditions and consequences depending on actions. However smart contracts are completely digitalized in blockchain. The importance of smart contracts comes in managing interactions between nodes and participants of the system based on data. Smart contracts like any other node have addresses in the blockchain. Triggering a smart contract is done by addressing a transaction to it.

3) Mining: Mining is one of the most important processes in blockchain technology. Mining is the act of validating new blocks so that they could be added to the blockchain. Mining is done by providing ”proof of work” to validate a block, i.e, each block contains a mathematical puzzle that needs to be solved by the miner in order to provide proof of work. This process can be very expensive regarding computational power. Miners are individuals or organizations having dedicated considerable computational power for mining and maintaining the blockchain.

B. Previous Works

After a large investigation in this field, we found that blockchain is a great candidate for future decentralized IoT architectures and models. The authors in [1] define a complete architecture of three layers, where blockchain is used as the storage layer. The authors also define a data management and a data sharing protocol, along with a study on different mechanism and their impact: Direct blockchain access, Server

Client access, Publisher Subscriber access. Based on this study, we decided to use publisher subscriber mechanism in our solution. Blockchain is also considered in Healthcare. Authors of reference [4] define the general role and future of blockchain technology in healthcare. In [5], a smart contracts based solution is defined by the authors, to manage the access to electronic medical records. Contribution in [5] is very important since it defines the basic use of smart contracts in managing relationships between different parties of the blockchain. Another study that focuses on the use of smart contracts in IoT context [6] presents important aspects and point of views regarding smart contracts. A closer domain to the Internet of things is having medical data generated by medical sensors, and Wireless body area network (WBAN), an approach that is discussed in [7]. Authors use a cryptographic approach applied to the blockchain, and an intermediate device to manage data flow and connect to the blockchain. [8] describes also a cryptographic approach where blockchain is implemented as a solution to protect personal data. Based on these results, studies are moving towards user-centric solutions. Our research aims to take into account all these ideas to derive a well defined solution that will be discussed later on in this paper. In the next section we will define the problems with applying blockchain as solution in an IoT context.

III. The Challenges of Blockchain Technology with IoT

Like already mentioned, blockchain is a peer-to-peer decentralized technology, it provides transparency, and gets rid of the need for third parties. With these advantages comes some important issues. First of all, in the nature of blockchain, it is a replicated ledger on all nodes connected to the blockchain. Therefore storage might not be very effective if it is being done on the blockchain itself. This imposes an issue for IoT devices such as the typical and normal sensors. On the other hand, since we got rid of third parties, information added to the blockchain means blocks are added and need to be validated: Mining. Mining is an important process that should be taken into consideration, because it demands computational power [9]. Another issue is scalability; when the number of nodes increases, the number of transactions increases and the mining and validation process takes longer to be completed. In our paper, we aim to address specifically the mining problem, and the time taken to complete transactions from end to end. We focus on the factors and parameters that might affect the mining process such as the block size and the block time. In the next section, we discuss our proposed model and the proposed solution.

IV. Model of the System

An interesting approach for solving the problem of the high computational power needed in order to communicate with the Blockchain, is to introduce a centralized-decentralized combined architecture i.e introducing intermediate servers between IoT devices and the Blockchain. A very good candidate for such an intermediate is a cloud server, easily accessible by clients, and can handle high computational power. Another effective approach is the use of a Publisher-Subscriber data sharing mechanism. The authors of reference [1] show us that a Publisher-Subscriber mechanism is very efficient in terms of filtering Data, and introducing intermediates to handle electrical power and computational power consuming protocols and technologies such as the Blockchain. On the other hand, our model is also based on Smart Contracts. Smart contracts allow us to maintain rules, authentication, and communication between different nodes and parties of the system. The final part of the model is the use of an off-chain database, since the transaction of messages containing data of important sizes directly decreases the performance of the Blockchain, and in the nature of this technology which is replication on every single node, the storage of files and data on-chain will cause delays and worse performance. One candidate for off-chain database is IPFS [10].

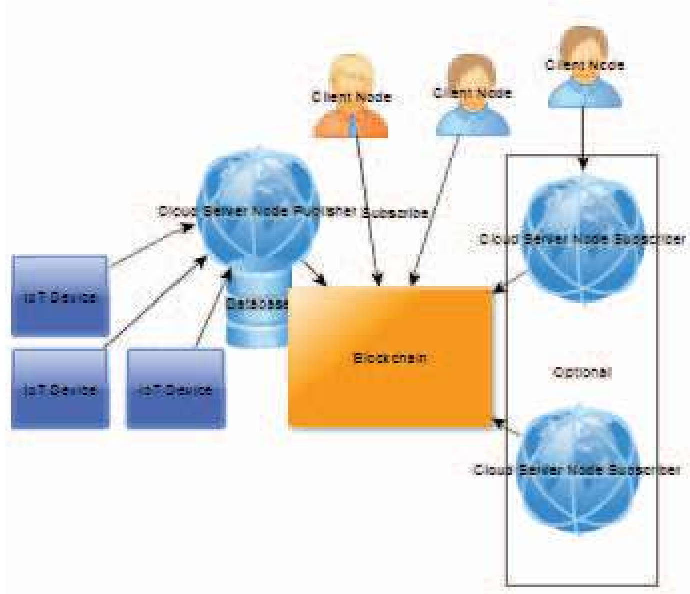

Fig. 1. Proposed architecture, IoT devices connected to the Publisher node which is connected to the blockchain, clients either connecting to subscriber nodes or directly to the blockchain

We can describe our model by first defining some notations that will be later used in this paper:

- P_k: Publisher number k
- d_i^k: Device i connected to publisher k
- S: Subscriber
- $store()$: Function to store in the off-chain database
- $Gen()$: Function to generate pointer to location in the database
- Eth: Client
- H: The Hash pointing to the location in the database
- $Sub(P)$: Function that determines if a certain subscriber is subscribed to a certain publisher

Figure 1. describes our proposed architecture in a general matter. There are two study steps in this model; the first step is to simplify this architecture by working on the blockchain re-

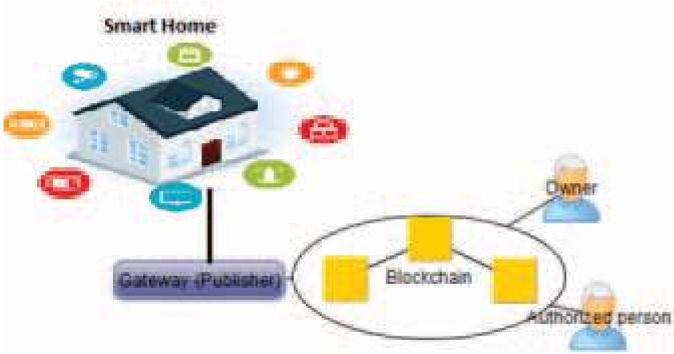

Fig. 2. Architecture for smart home scenario, including the gateway to the blockchain, the owners and authorized people to access the data

lated part only, thus the communication between the publisher nodes and subscriber nodes regardless of the users connected to it. The second step is to focus on the end user part, its connection to the subscribers, and the IoT devices and how they are connected to publishers, thus how data is being sent and organized. In the next section, we will go into details of the main components of the system, as well as the proposed solution including the protocol and the smart contracts design.

V. Proposed Solution

To design our solution, we have first specified a potential real life scenario. In this section, we will first define the scenario, then define the Smart contracts model, and finally detail the access protocol on publishers and subscribers.

A. Defining the Scenario

One of the major project today's are "Smart Cities", "Smart Buildings", and "Smart Homes". How can a person monitor the devices in his house with the complete knowledge that he is the only one that could access that data? Therefore, in order to elaborate our solution, we propose the following scenario: a smart building where sensors on different objects are connected, always online, and users need to monitor the sensors at every moment. The idea is to secure access to the data generated by these sensors. This can be done by protecting the information about the physical location of the data. As described in the previous section, the main components of our system are : the Blockchain, Smart Contracts, Intermediate servers and an Off-chain Database. The main parties of our system are: Publishers and subscriber. The main end parties are: IoT devices and users. Figure 2 presents the proposed solution architecture based on the general architecture presented in figure 1.

The idea highlighted figure 2 is based on the fact that normal sensors in a smart building cannot directly be connected to the blockchain. Thus, to reach a certain level of peer to peer communication, we use a gateway that connects the Smart Home to the blockchain. This gateway will play the role of the publisher, since it will publish the data from different sensors in the smart home. Owners are the people who own the smart home, thus have access normally to the data being generated.

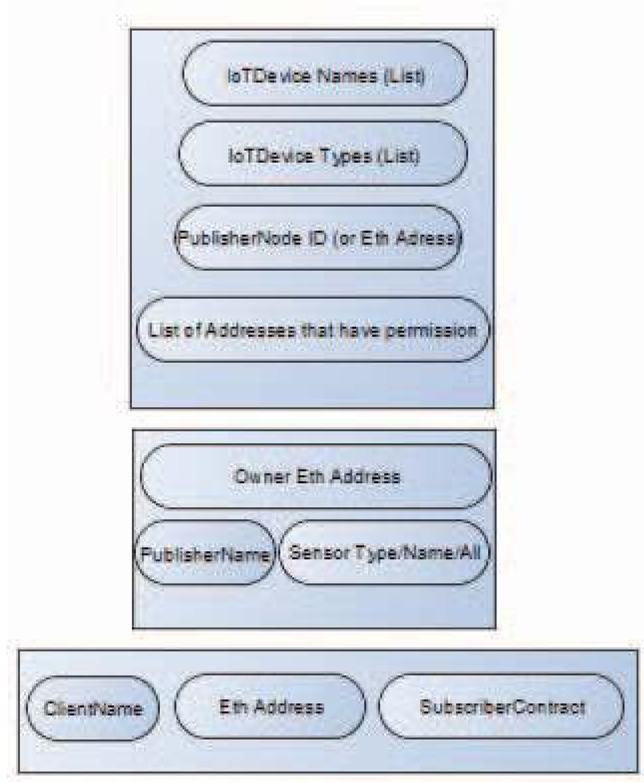

Fig. 3. The Smart Contracts used in our solution; Publisher Contract, Subscriber Contract and Client Contract

Authorized people are those who are granted access by the owner.

B. Defining the Smart Contracts

In order to realize a successful implementation of Blockchain technology in the Internet of things context, Smart Contracts should be the corner stone of the system.

Figure 3 defines the structure of the Smart Contracts in our solution.

1) The Publisher Contract: : The first contract is the publisher contract, it will be deployed by the publisher. One can think about contracts as functions, that take input, checks it, and gives us a result. When a user wants to connect his smart home to the blockchain, he needs to define first his Gateway or publisher contract, giving an ID mapped to its Address (Ethereum [11] address in the case of ethreum blockchain), the list of IoT devices by name for ease of access and comprehension, and by type to manage the publisher subscriber data sharing mechanism. And finally, the list of addresses that have permission, usually these are the addresses of subscribers, who are the authorized persons chosen by the owner himself.

2) The Subscriber Contract: : The second contract is the subscriber contract, that is normally deployed by the owner. The subscriber contract must contain the address of the owner in the blockchain, and the publisher or publishers that he is subscribed to, along with the Sensors he chose to receive data from. The sensors can be chosen by type, by name or all the sensors connected to a particular publisher. This is the critical component of the publisher subscriber algorithm, since based on information stored in that contract, data generated will be filtered at each receiving node.

3) The Client Contract: : The third contract is the client contract. It serves as a mapping contract between normal nodes, or clients connected to the blockchain, and their respective subscriber contracts. It contains the client name, so that it would be simpler for clients to communicate between each other using a frontend application. This name is mapped to its address in the blockchain. It also has the subscriber contract belonging to the client.

C. The Protocol Description

After describing the contracts, this part describes how the publisher node acts upon receiving new data generated by an IoT device connected to it. The steps that should occur are (1) storing the data in the off-chain database (IPFS in our case), (2) generate the hash pointing to the location in the database, then (3) sending that hash to every node on the blockchain that is subscribed to this publisher. The transaction are sent through the blockchain client, which is an ethereum client in our case.

Publisher node upon IoT device generating new data:
$Da = Data(d_i^k)$
$if(\exists a)$
    $store(Da)$
    $H_P = Gen(HashDa)$
    $foreach(S_j / S_j \in sub(P_k))$
        $Eth.send(H_P, S_j)$

The second part describes how the subscriber should act upon receiving a transaction sent by a publisher it is connected to. The ethereum client should receive a transaction from the publisher, containing the hash. The subscriber then tests the address that sent that transaction, in order to counter any attempts of spam. When it is done, there are two approaches; either the client is directly connected to the blockchain (the client in this case deploys his own subscriber), so he directly receives the transaction through his ethereum client, or he is connected to a certain subscriber. This part implements the idea of having a subscriber handling the data flow, by broadcasting the received message containing the hash to all connected clients.

Subscriber upon receiving new data from publisher
$H_s = Eth.rec()$
$if(\exists H_s \&\& Eth.address(sender) == Eth.address(P_k))$
    $foreach(C_j)$
        $send(H_s, C_j)$

The third part simply describes the fetching of the data using the received hash in the previous part. He connects to the IPFS node in our case, and uses the hash to fetch the data generated by the sensor.

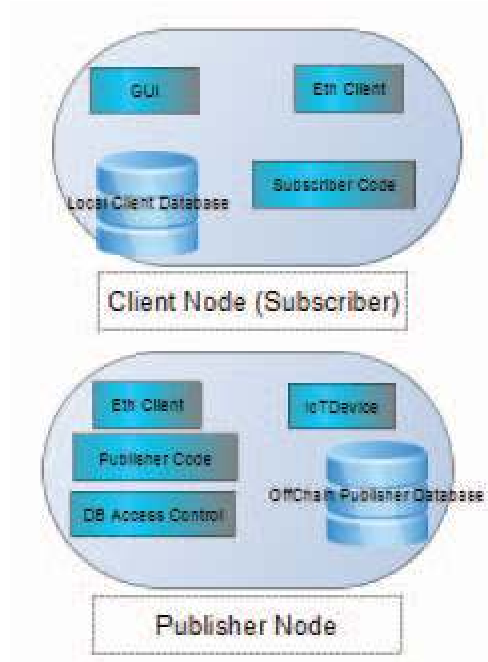

Fig. 4. Client Node and Publisher Node along with the components of each node

Client Data Access
$H_s = rec(S)//$ will allow client to determine the database based on the address DB_P
$if(\exists H_s)$
    $C.connect(DB_P)$
    $C.fetch(H_s)//Data = C.fetch(H_s)$

In figure 4, we can observe each part of the subscriber and publisher nodes.For the subscriber node, the Graphical user interface (GUI) interacts with the user, the eth client connects the node to the blockchain, and the local database stores the data received, it works like a cache for fast access in the future. For the publisher node, eth client IoT devices are connected to the node, the publisher code manages the publisher subscriber mechanism in the publisher, the off-chain database contains all the data generated by IoT devices from this publisher and the Database Access control manages the access to the database. It is also important to notice one important step which is mining, since with every transaction taking place, nodes on the blockchain (miners) need to validate the transaction which will make it longer for the message to be delivered from publisher to subscriber.

VI. Implementation

As mentioned in the previous section, mining is an important factor that should be taken into consideration. To optimize the performance of such a system, many factors should be studied, and many constraints are presented. To optimize our solution, the time for a transaction to take place should be

minimized. In fact, the transaction time depends on the rate and size of the data; $t = f(r,s)$. The rate is important because it affects the mining process, if rate generated by IoT devices increases, transaction time will increase. The mining process depends on the capability of the device processor to validate that transaction, thus the time depends also on it. The best way is to lower the rate and size of the data, however, constraints exist. We can't control the CPU power, and we can't control the rate and size of data from some sensors. For example, an ECG sensor is critical, and has to send data more frequently than a temperature sensor. Our approach is based on the mining time. N_d Number of devices connected to a Publisher r_d : Rate of new data generation, thus new blocks in the blockchain The total data rate generated by one publisher that has N_d connected devices,assuming that all the devices have the same rate r_d is $N_d \times r_d$ new data per second.
New Data $\Rightarrow$ new transactions $\Rightarrow$ new blocks to be mined. The mining rate depends on the block size N_t and the time to mine one block T_b. The block size depends on the number of transactions that one block can fit, it varies from One blockchain to another; Bitcoin fix size,Ethereum Dynamic So the mining rate is $\frac{N_t}{T_b}$
Case 1: Max Delay $=T_b$ if $N_d \times r_d \leq \frac{N_t}{T_b}$
Case 2: Max Delay = $N_b \times T_b$ where N_b is the number of blocks to fit all transactions if $N_d \times r_d \geq \frac{N_t}{T_b}$

N_b can be calculated supposing t=1s $N_b = \frac{N_d \times r_d}{N_t}$
In the same way this can be generalized on data generated by all publishers to obtain $\sum_{n=1}^{N_P} N_d^P i \times r_d^P i \geq \frac{N_t}{T_b}$

This study aims to calculate the time needed for a transaction to be mined and a bloc to be validated, without taking into consideration the work and time taken by the subscriber and publisher mechanism.
To implement our solution, we have chosen to use a framework called Embark [11]. Embark is based on the idea of Decentralized applications (DApps). It implements ethereum blockchain, IPFS database, and Whisper protocol used to send messages between multiple DApps. We have chosen to work with the go ethereum (geth) client, along with SolidityC [12] language for smart contracts programming, HTML, Javascript and JQuery for frontend and GUI programming. We have successfully created our private blockchain, with a couple of nodes, and notification smart contracts upon new data updates. We have also successfully connected a Raspberry Pi to our Blockchain, which will work as the gateway for the system. We have two gateways so far, a laptop, and a Raspberry Pi. The latter helps simulate real data from real sensors for future studies. In the next section, we will discuss our future work. Our goal in the future is to implement all the components in Figure 4, and complete our architecture, along with the protocol, in order to study and evaluate the performance of this model when it increases in scale. In fact, the real problem lies with having an important number of IoT devices, generating data simultaneously.

VII. Conclusion and Future Works

The objective of this work was to consider the possibility to use blockchain technology in the area of security in IoT. We have presented an architecture solution designed for that purpose that is based on contract model between a provider and a consumer of data. To cope with the size of the data to store, we have proposed to associate the blockchain with complementary off-chain database technology. The block contains the main contract information as well as reference to where the complete data is stored. We have also presented an analytic model to study the performances of the mining process that has a high impact on the overall response time of the system. In our future work, we plan to go further in the implementation of the system and perform large scale tests with physical and virtual IoT devices.

Acknowledgments

This research was supported by The University of Evry Val dEssonne and partially by the Lebanese University. This work was conducted in the frame of the PHC CEDRE Project N37319SK. We also thank our colleague Dr Massum Hasan for the early discussions on the topic.

References

[1] Sayed Hadi Hashemi, Faraz Faghri, Paul Rauschy and Roy H Campbell, ”World of Empowered IoT Users”, *2016 IEEE First International Conference on Internet-of-Things Design and Implementation (IoTDI)*, 2016

[2] S. Nakamoto. (2008). Bitcoin: *A Peer-to-Peer Electronic Cash System.* [Online]. Available: https://bitcoin.org/bitcoin.pdf

[3] Nick Szabo, *The Idea of Smart Contracts*, [Online], Available: fon.hum.uva.nl/rob/Courses/InformationInSpeech/

[4] Matthias Mettler, M.A. HSG Boydak, ”Blockchain Technology in Healthcare The Revolution Starts Here”, IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom), 2016

[5] Asaph Azaria, Ariel Ekblaw, Thiago Vieira and Andrew Lippman, ”MedRec: Using Blockchain for Medical Data Access and Permission Management”, , International Conference on Open and Big Data (OBD), 2016

[6] Konstantinos ChrisTidis and Micheal Devetsikiotis, ”Blockchains and Smart Contracts for the Internet of Things”, *IEEE Access (Volume: 4)*, 2016.

[7] Jie Zhang, Nian Xue, and Xin Huang, ”A Secure System For Pervasive Social Network-Based Healthcare”, IEEE Access (Volume: 4), 2016

[8] Zyskind, Guy, and Oz Nathan. ”Decentralizing privacy: Using blockchain to protect personal data.” In Security and Privacy Workshops (SPW), (2015) IEEE, pp. 180-184.

[9] Mining in Bitcoin, [Online], Available: https://en.bitcoin.it/wiki/Mining

[10] IPFS - Content Addressed, Versioned, P2P File System, [Online], Available: https://ipfs.io/docs/

[11] Ethereum Whitepaper, [Online] Available: github.com/ethereum/wiki/wiki/White-Paper

[12] Embark Framework, [Online], Available: https://github.com/iurimatias/embark-framework

[13] Solidity Introduction to Smart Contracts, [Online], Available: http://solidity.readthedocs.io/en/develop/introduction-to-smart-contracts.html

[14] Kyle Croman, Christian Decker, Ittay Eyal, Adem Efe Gencer, Ari Juels, Ahmed Kosba, Andrew Miller, Prateek Saxena, Elaine Shi1, Emin Gun Sirer, Dawn Song, and Roger Wattenhofer, ”On Scaling Decentralized Blockchains”, *Initiative for CryptoCurrencies and Contracts (IC3)*, 2016