Making Case for Using RAFT in Healthcare Through Hyperledger Fabric

Anastasios Alexandridis
ECE. McGill University
Montreal, Canada
anastasios.alexandridis@mail.mcgill.ca

Ghassan Al-Sumaidae
ECE. McGill University
Montreal, Canada
ghassan.al-sumaidae@mail.mcgill.ca

Rami Alkhudary
LARGEPA. Université Paris II Panthéon-Assas
Paris, France
rami.alkhudary@u-paris2.fr

Zeljko Zilic
ECE. McGill University
Montreal, Canada
zeljko.zilic@mcgill.ca

Abstract—Blockchain technology is enabled by consensus algorithms to manage the relationships among several economic or business operators without human intervention. With the help of consensus algorithms, distributed systems can reliably reach agreement even if part of the system is faulty. Blockchain yields many benefits, among others, traceability, transparency, and security. We consider using the RAFT consensus algorithm to achieve robust and scalable decentralized applications, with focus on healthcare. We propose a stylized healthcare network, enabled by RAFT and built upon Hyperledger Fabric to showcase the use of RAFT in healthcare blockchain. However, RAFT is by no means limited to healthcare record systems, and can be applied to any other record system and value chain. Our paper offers several insights to those working in value chains and information management-related fields. In addition, we end our study with some future research avenues that may inspire managers and scholars to build or refine new decentralized systems in healthcare and other related fields.

Index Terms—blockchain, RAFT, consensus algorithm, Hyperledger Fabric, value chain, healthcare

I. INTRODUCTION

The RAFT consensus algorithm [1] produces agreement on the log of activities on the replicated state machine. While RAFT can be applied to numerous applications, we are interested in its application in blockchain. Blockchain arose as the peer-to-peer electronic cash system Bitcoin proposed by Nakamoto [2] (we still do not know the real identity). Bitcoin has solved the problem of double spending on the internet, allowing a value exchange of digital assets or financial transactions without intermediaries or trusted third parties’ intervention. The underlying blockchain technology relies on three relatively old technologies: distributed ledger technologies, cryptography, and consensus algorithms.

Blockchain can be defined as a distributed ledger or decentralized database of digital records chronologically organized and immutably registered following a precise consensus mechanism [3]–[5]. Consensus is necessary in decentralized databases with no trusted third parties to ensure reliable data. Blockchain operation can be better understood by explaining the main features of the Nakamoto model (public blockchain).

First, each blockchain network must have a consensus mechanism, often translated into the mining process, to verify, link, and securely register transactions or data records by cryptographic proofs. Second, in theory, the network is fully decentralized since each user can have an identical copy of the blockchain ledger, maintain its integrity, and play an essential role in the mining process. Third, transactions or digital records become traceable, transparent, and immutable when registered, although users are anonymous in the public blockchain network [3], [6].

Nakamoto’s proposition was mainly limited to cryptocurrencies and distributed exchange systems. Many years later, research on blockchain in record systems such as healthcare records extended in all directions [7], particularly in the grey literature. Although the public blockchain model (Nakamoto model) was proved to add value in a few healthcare applications [8]–[10], the current mainstream is using the private blockchain model [11]–[15]. The private blockchain model is different in that some economic or business operators (I) decide to work together building a limited access (closed) network, and (II) define the mechanism of the consensus that has to be reached to validate transactions as there is no need for mining [16]–[18].

PAXOS [19] is one of the very first attempts to achieve scalable and fault-tolerant consensus between multiple distributed systems. This mechanism is based on two principles by which consensus can be achieved across distributed systems. The first principle is the existence of a node that proposes a certain value so that all other nodes in the network adopt that certain value, called the proposer node. The proposer should associate this value with a unique number, so that the other nodes will commit to this value if they have never seen a number higher than this number before. The second principle is that there is a type of nodes making a decision about the value sent by the proposer, a so called acceptor. The acceptors either confirm the value and send the approval back to the proposer, or they partially accept it by guaranteeing that none of the other values received from different proposers are accepted, or they reject the value altogether and inform the
proposer of this rejection. However, PAXOS was not as easy to put into practice for many reasons, including the lack of tools for handling physical failures in distributed manner. It is hard, for instance, to treat a plethora of disk corruption mechanisms under data replications among multiple actors, as it result in an unknown number of corner cases and with uncertain correctness provability, especially when malicious actions enter the play. These difficulties in understanding and implementation have cast a shadow on applying PAXOS in the real world, so extra effort was required to achieve real consensus. For this reason, RAFT [1] was proposed as an alternative to the functions of PAXOS, but one that is much easier to implement and understand. In a nutshell, RAFT is a leader-based mechanism, which means that data can only flow from the leader node to the other server nodes in the network. Consensus in this mechanism can be achieved mainly through three processes: immediate election of a network leader in case of failure of the current leader, replication of the log entry across all nodes, and permanent safety of the committed log entry to persistent storage. More details on RAFT consensus are provided in Section II after we introduce our application domain.

Blockchain yields many benefits in healthcare and other related fields handling large amounts of data sets, among others, traceability [20]–[22], transparency [23]–[25], and security [26]–[28]. The literature explains many consensus algorithms that can be used in decentralized healthcare applications, which are the focus of this paper, for public and private blockchain systems. For example, Yazdinejad et al. [8] proposed a blockchain configuration for a decentralized hospital network enabled by Proof-of-Work. Similarly, Yang et al. [29] designed a medical data sharing system based on Proof-of-Stake. Wang et al. [30] proposed an intelligent healthcare supply chain based on Delegated-Proof-of-Stake. Zhu et al. [12] introduced a consortium blockchain system using Proof-of-Authority. Zghaibeh et al. [11] proposed a multi-layer and intelligent health management system based on Practical-Byzantine-Fault-Tolerance. To end with, Wang et al. [31], [32] proposed a blockchain-based eHealthcare system interoperating with WBANs, with a multistage approach to increase performance and throughput, that proposes a leader and follower node configuration of a Crash-Fault-Tolerance implementation that uses Apache Kafka.

In trying to apply RAFT to blockchain, the literature falls short in explaining how the RAFT consensus (equivalent to PAXOS in fault-tolerance and performance) adds value to decentralized healthcare applications. In response, we propose a stylized healthcare network enabled by RAFT using Hyperledger Fabric as a proof of concept of its applicability. Our paper is hoped to offer insights to providing value chains in information management-related fields. In addition, we end our study with some future research avenues that may inspire managers and scholars to build or refine new decentralized systems in healthcare and other related fields.

II. RAFT CONSENSUS

With the recent increase in data breaches, it is becoming increasingly urgent to seriously consider a protection technique that ensures the security of data and guarantees seamless access when needed. Replication of data on distributed machines has been proposed in the literature as a solution for this dilemma. However, whenever the term "distributed systems" is mentioned, the first question that arises is how these systems achieve consensus on the data state. Even in the case that consensus is reached, the question is how can it be guaranteed that all nodes are updated synchronously. These requirements are very urgent for a fault-tolerant distributed system that can function despite the failure of some nodes. Replicated And Fault-Tolerant (RAFT) [1] is a state machine replication protocol proposed to meet these requirements and solve the problem of inconsistency of distributed databases [33].

Nodes implementing RAFT can take one of three roles [34]: Leader, Candidate, and Follower. The leader node is responsible for interacting with clients on the network. Any client’s request received by the leader is directly forwarded to the followers for confirmation. The followers are responsible for acknowledging the requests and sending them back to the leader node. When a client attempts to contact the followers directly, they forward the request to the leader node. All nodes are initiated as follower nodes in their very first inception. The network should have only one leader node at a time. If this leader node fails for any reason, it paves the way for any follower node to become a leader node. Fig. 1 shows the transitions of the node states.

In RAFT, time is divided into periods of arbitrary length, numbered consecutively. The election process to select a leader usually begins at the start of each period and ends the moment a winning node is elected [35]. However, the election can also be conducted if no leader node has appeared in a given period. The nodes in RAFT communicate through Remote Procedure Calls (RPCs), and consensus can be reached through two types of RPCs: RequestVote and AppendEntries. The leader node periodically sends a message to all followers to maintain its authority. There are two types of AppendEntries messages that the leader node can send to the followers; one is used to replicate the log entries when a new request is received.
from the client. The other one does not contain any log entries and is only used to inform the followers about the continued validity of the leader node. However, followers have the right to change their status to candidate once a certain amount of time has passed without hearing from the leader.

Once a follower does this, it starts voting for itself and sends a RequestVote to all other nodes. Then, three scenarios can occur. The candidate node could get the most votes from the nodes and then declare itself the winner of the leader position in the network. The candidate node does not win this contest because another node has declared itself the leader. Last but not least, the time allotted for the election has elapsed without a winner. In this case, the call for the election is restarted. After the winning node is declared, it sends a message to all nodes to inform them of its authority and block the way for a new election for the rest of the term. As mentioned earlier, the new leader then begins processing the requests that come in from the clients. Whenever a new request arrives, the leader sends an AppendEntries message to the followers and waits for their confirmation that this request has been successfully replicated. If the followers do not respond appropriately, the leader sends RPC messages until the log is updated on all follower nodes.

The fork problem is prevented in the current Hyperledger Fabric enabled by RAFT since all network members are well authenticated. Thus, computationally complex consensus mechanisms (mining) are not necessary. Moreover, RAFT is efficient and implementable in private value chain networks [36], [37].

III. STYLIZED HEALTHCARE NETWORK BUILT UPON HYPERLEDGER FABRIC

The example of a hospital network can be used to present our stylized healthcare configuration. A hospital or group of hospitals will typically maintain some kind of healthcare network, which facilitates communication between employees and machines, as well as a database that includes, but is not limited to, patient data such as examinations, prescriptions, or other information. It is critical that all this information exchange and storage is done in a secure and private fashion, both inside the hospitals as well as other external organizations such as laboratories, pharmacies, insurance companies, or other medical services. Other organizations are not in the scope of this work, and further work would need to examine their participation in the healthcare network.

In order to make use of blockchain, the health network can be designed as a blockchain network. As mentioned, a group of hospitals essentially has a secure communication network, and a database where data storage takes place. Different entities can interact securely with the database. Every hospital has an unspecified number of end users; employees, such as doctors, and machines, such as biomedical sensors, that can interact with each other and the database. The hospitals can also interact with each other. This is necessary in case patient data needs to be transferred to a different unit or even hospital.

Every hospital has back-end servers, which function as the blockchain database, or ledger. These servers are a core part of the blockchain network, and are hosting blockchain peer nodes who maintain the ledger, and ordering nodes. The servers can be physically separated into multiple machines, such that if there is failure in one machine, only one node will be compromised. Further redundancy can also be achieved on a hardware level; for instance the disk of a peer node which maintains a copy of the blockchain ledger can be redundantly mirrored, ensuring the peer node can continue working even if there is damage on the physical hardware. All the hospital back-end servers are connected together, and part of the same blockchain network. This can achieve data redundancy and replication, in essence producing a distributed ledger both in terms of machines, but also from a geographic point of view, and providing shared data access to all the hospitals.

The end users that interact with the network are the front-end, and can be viewed as clients of the system. Through the use of a client application external to the blockchain, the end users can interact with the blockchain and execute smart contracts on the peers. For example, if a doctor needs to insert some information about a patient in the ledger, they would use an application to run a smart contract on the peer, and provide the information as a parameter. From the doctor's point of view, the information is stored in the ledger as if it were a traditional database system, however the process is a little more nuanced than that. The client application will submit a transaction proposal to the blockchain network, which needs to be endorsed by a predefined number of peers, then ordered into a block by the ordering nodes, and then validated, before it is finalized in the ledger. During this process, the peers and ordering nodes check against policy criteria, validate that endorsements came from appropriate entities, and check the version of the ledger to ensure data integrity is not compromised, by repeatedly signing, verifying, and authenticating payloads as they pass through the network. All this process is abstracted from the end user.

Fig. 2 illustrates the process of querying or updating the ledger. When a doctor needs to retrieve or insert information in a patient health record, they can use application (A) to do so. A will connect to peer 1 (P1) and execute the appropriate section found in smart contract (S1) by submitting a transaction proposal as discussed above. It is worth mentioning that for a simple ledger query and information retrieval, P1 will endorse the transaction proposal and return the information from ledger (L1) to A in the form of a proposal response. No transaction will be committed to L1.

However, if a ledger update is needed, there are additional steps as discussed earlier. After being endorsed and a response is sent to the application, the transaction goes to orderer 1 (O1) for ordering, which will order this and other transactions from other applications into a block, that is then sent to all peers in the network in order to be added to the ledger. Fig. 3 goes into more detail about the transaction flow. Here, application A1 has received endorsement (E1) for its transaction (T1), and forwards it to O1 which orders it into a block. O1 has received more transactions, such as T2 from application A2, which in fact arrived earlier. Once a block is formed, it is sent to all
peers in the network as shown in Fig. 4, which will validate it before appending it to the ledger, finalizing the process.

For all intents and purposes, necessary components of the system such as the Membership Service Provider (MSP) and Certificate Authority (CA) are abstracted here, however they are very necessary. Authentication of the different actors in the network constantly takes place, and is required to determine the exact permission each one has over different resources and information. For the purpose of this stylized healthcare network, an example of a group of five hospitals can be considered. These hospitals can be viewed as one organization, for example an organization named "hospitals of a city". Each hospital can have two servers, with one server being a peer node, and another server being an orderer node. All nodes are members of one channel. The peer nodes have smart contracts, also called chaincode, installed on them. As mentioned earlier, an end user with the right permissions that come from their identity can execute smart contracts. More particularly, a doctor will have permission over reading and writing data concerning their own patients, whereas a hospital supervisor may have read permission to all patient data. A system administrator will have full access. Access control through identity verification can be used to ensure that depending on the role of a system user, the appropriate permission is allowed.

In practice, the transaction flow is as follows. In the start, an end user, such as a doctor, needs to have a certificate issued by the CA in order to authenticate in the network. The doctor will use an application to initiate a database, or ledger, update, for example updating a patient’s health record with a recent test results, which will be sent to the peer node found in the hospital. It should be mentioned that the application the doctor is using could communicate with another peer located remotely, and it is not necessary to contact the local peer, or only one peer. The smart contract installed in the peer will, among others, contain a function for the doctor to submit the most recent patient test results. Additionally, an endorsement policy in the smart contract will specify how many peers are needed, in this case it can be assumed that one is sufficient.

The ledger update will initiate by the application initiating a transaction. This update will trigger the function in the smart contract that updates the patient test results. It will also contain certain parameters, for instance who the patient is and the actual test results. The certificate of the doctor is also used in this stage to sign the proposal.

The proposal will then go to the peer which will endorse it by verifying the signature, as well as that the proposal is unique and correctly formatted. This can ensure that it comes from a proper entity and not, for example, an attacker communicating with the network or a malicious application not working properly. Signature verification is done using the CA and MSP mentioned above. In short, the former ensures the certificate is valid and verifies the identity while the latter ensures that the particular identity has the right permissions, however, this is outside of the scope of this work. The endorsing peer will then execute the transaction without yet updating the ledger with the new patient test results, update the transaction with the results, sign it, and send it back to the application as a proposal response. The application will then verify the endorsing peer signature and correctness of the response, as well as compare and merge responses together in the case of multiple endorsing peers, which is not the case here. As discussed earlier, if the doctor was simply reading the ledger, the transaction would not be sent for ordering or committed to the ledger, however the patient’s recent test results need to be inserted in the ledger. For this reason, after receiving the proposal the application will ensure the endorsement policy holds true, in this case that one peer has endorsed it, and then send it to the ordering nodes.

The ordering nodes will order this and other transactions chronologically and package them into blocks, which are saved and distributed to all the peers in the network. It must be mentioned here that ordering is deterministic and not probabilistic, that is, the blocks are final and there cannot be a ledger fork. Last but not least, the peers that receive the block will need to validate all the transactions in it, to ensure the endorsements are correct and match the endorsement policy.
as well as no transaction is invalid. In the latter case, those transactions are marked as invalid and do not update the ledger. In the end, the peers will immutably append the block to the blockchain, and all the ledger updates mentioned in all the transactions inside the block will be performed on the ledger, updating its state to be the most recent one.

As discussed, attacks are eliminated by verifying the signature of the end user. Because the end user applications are external to the blockchain, anyone who can replicate one and forge a communication packet that for instance includes a transaction proposal would in theory be able to interact with the network. However this assumes that they are able to replicate the transaction proposal correctly, otherwise it would be rejected. Further, the lack of verification would prevent an attack from continuing, and it would be recorded for further investigation in the incident. Even in the case of identity, or in practice certificate, theft which is unlikely, the attacker will need to format a proposal correctly, and only have limited access to the network, as permissions are allocated only to those who need them. This shows that the system is secure, but of course not impenetrable, however by combining multi factor authentication technologies like biometrics or physical tokens, the user error can be virtually eliminated.

IV. THE ROLE OF RAFT IN OUR PROPOSED CONFIGURATION

In this section, we make a case for using RAFT in our stylized healthcare configuration that is built upon Hyperledger Fabric discussed in the previous section.

As mentioned, an organization serves as an umbrella for five different hospitals (H1, H2, H3, H4 and H5). Our hospitals have 10 servers, and each hospital has a peer node P and an orderer node O. Since only the orderer nodes are performing consensus and thus eligible to be a leader node, the organization needs at least (O/2)+1 as a quorum, which must endorse the request sent by the leader. More clearly, there is a need for at least 3 majority nodes and the organization can be tolerated with 2 failed nodes. Orderer nodes O1 → O5 can be in one of three states: follower F, candidate C, and leader L. It can now be assumed that they are in their initial phase of operation, and the first cluster starts up (term 1). Now, each node among the orderer nodes O1 → O5 will pick a different random timeout and then start counting it down. This timeout represents the interval that this node will wait until it moves from its F state to C state. When the time runs out for each node respectively, the following scenario will happen:

\[
\begin{align*}
O_1H_1 & \xrightarrow{\text{vote request}} O_2H_2, O_3H_3, O_4H_4, O_5H_5 \\
O_2H_2 & \xrightarrow{\text{vote request}} O_1H_1, O_3H_3, O_4H_4, O_5H_5 \\
O_3H_3 & \xrightarrow{\text{vote request}} O_1H_1, O_2H_2, O_4H_4, O_5H_5 \\
O_4H_4 & \xrightarrow{\text{vote request}} O_1H_1, O_2H_2, O_3H_3, O_5H_5 \\
O_5H_5 & \xrightarrow{\text{vote request}} O_1H_1, O_2H_2, O_3H_3, O_4H_4
\end{align*}
\]

Let us suppose that O5H5 runs out its time out first. In this case, O5H5 will be getting into C state and then increasing its current term by 1. Since O5H5 is now eligible for a leadership position, the very first step in this direction is to vote for itself and then send a RequestVote to all other orderer nodes O1 → O4. These orderer nodes must now receive the request, and if these nodes have not yet cast their vote, they reset their term to match the term of O5H5 and then cast their votes to O5H5. Has the majority responded to the vote request, O5H5 is now the leader of the entire organization. The leader O5H5 then begins to send AppendEntries messages to all orderer nodes O1 → O4 from time to time to inform them of its continued authority as a leader, and to prevent any new election during the current term. All orderer nodes in all hospitals must then reply to the request message sent by the leader. However, the current term could be terminated the moment that O1 → O4 nodes are no longer hearing from the O5H5 node. In this case, O1 → O4 will assume a node termination scenario and switch to state C. Fig. 5 gives an overview of the proposed scenario.

As indicated above, the organization is now led by the O5 server located in hospital 5. Any ledger updates in the organization should now go through O5. In the previous example, a doctor needed to update a patient’s health record with recent test results. It can be assumed here that this doctor is working on H3 and interacts with P3, even though the latter will not always be the case. After the transaction proposal has been endorsed and a response is sent, the application will submit the transaction to O5 where it is arrives among other transactions in need for ordering. However, even though these transaction have now been received, the state of O5 will not be updated, given that the other orderer nodes have not yet acknowledged the submitted transactions. In order for the transactions to be packaged into a block, O5 must now replicate them over O1, O2, O3 and O4. If it is assumed that O1, O2, O3 and O4 successfully added the transactions to their logs, while O2 has failed to synchronize, then in this case the transactions will still be committed on O5 based on the majority’s confirmation. Finally, O5 updates its log, creates the block and distributes it to all the peers.

It can be shown that the resulting blockchain transaction log is kept consistent due to RAFT safety guarantees and that the blockchain will be made correct and available for arbitrary sequences of transactions. Closer inspection of the leader election in RAFT shows that the algorithm maintains suitable leaders in a dynamically evolving situation.
V. Conclusion and Future Work

Our paper shows how RAFT can add value to data record systems and value chains, in particularly healthcare, using Hyperledger Fabric. This matters because the academic literature falls short in explaining how RAFT (one of the leading consensus mechanisms on Hyperledger Fabric) can play an essential role in decentralizing healthcare applications.

RAFT operates quickly and efficiently, on par with the PAXOS consensus. It can guarantee safety by using a new, two phase, mechanism when introducing new consensus nodes, thus avoiding the possibility that two majorities will be formed during configuration. It can also be seen that it is easier to understand and implement, speeding up both learning as well as development and maintenance of healthcare networks. Consequently, by accelerating blockchain development in healthcare, RAFT can provide decentralized healthcare applications that benefit from distributed ledger technologies. More precisely, powerful decentralized and append-only immutable ledgers can offer fault-tolerance in the form of redundancy, prevent cascaded failure, and offer trustworthy logging and auditing. Further, with the use of certificates and identity verification, strict access control and the property of non repudiation can be enforced. Other related fields and data record systems can also benefit from blockchain, and therefore RAFT consensus.

This paper proposes a healthcare network enabled by RAFT and built upon the Hyperledger Fabric architecture. From a technical perspective, we can confidently say that RAFT is a good candidate for healthcare and other related fields handling a large volume of data, and worth investigating further. In the future, the feasibility of implementation could be looked into further, as well as the inclusion of other organizations, such as the government, or insurance companies, for a complete system. Implementation and maintenance costs are also critical and need to be carefully examined if the system were to be built. Further improvements on the security and reliability of data are also possible. A system of this level of complexity has a multitude of attack vectors. While best practices can enhance security, it is critical to reducing user error and the impact of social engineering attacks. The use of biometrics and physical security tokens discussed above can be such an approach, eliminating the need for passwords. This can be investigated in more detail. Data reliability is also of importance, with smart contracts and endorsement policies able to offer some help. For instance, smart contracts can perform some basic error checking, and the more the system is automated and reliant on smart contracts, the more trustworthy the data can be. Finally, big data research can benefit from an append-only ledger found in a blockchain-based healthcare network. Hyperledger Fabric, in its recent versions, can take multi-thousands of transactions. However, we doubt that a decentralized healthcare system enabled by RAFT and built upon Hyperledger Fabric can handle a large quantity of data (in the case of implementing embedded medical sensors). We expect research regarding the scalability of decentralized healthcare systems that integrate big data analytics to be studied further, as blockchain and RAFT consensus are by no means limited to healthcare; there is value to be found in other related fields and distributed data record systems.

References


