Permissioned Blockchain based Architecture for Healthcare Data Management

MSc Research Project
Cloud Computing

Praveer Gupta
Student ID: 21143641

School of Computing
National College of Ireland

Supervisor: Rashid Mijumbi

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Praveer Gupta		
Student ID:	21143641		
Programme:	MSc. Cloud Computing	**Year:**	January 2022
Module:	Research Project		
Supervisor:	Rashid Mijumbi		
Submission Due Date:	15-Dec-22		
Project Title:	Permissioned Blockchain based Architecture for Healthcare Data Management		
Word Count:	7291	**Page Count:**	23

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: ……

Date: 15-Dec-22

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Permissioned Blockchain based Architecture for Healthcare Data Management

Praveer Gupta
X21143641

Abstract

A new era of "Smart Living" has begun all thanks to the Technology in Healthcare ecosystem and the ever-pervasive smart health devices. The idea of Smart Living has also fuelled the concept of Smart and "Healthy" living with the aim to increase the longevity and quality of one's life. Healthcare and Technology together aspires to raise the standard of healthcare by offering patients real-time services through the intelligent use of various gadgets and sensors. Since the healthcare ecosystem deals with such sensitive data it becomes imperative to secure the ecosystem, which otherwise, is not possible with traditional technologies. Permissioned Blockchain is one such technology to transform the world of Healthcare Data Management with its inherent properties of privacy and immutability. Hyperledger Fabric is one of the permissioned blockchain network which has a wide adaptability in healthcare use cases. The proposed architecture, this research work, focuses on how data is handled by various participating entities and keeps the control of the data with its owner, i.e., the patients. The data is managed and stored as Private Data Collections only at authorised Hospitals whereas the privacy and security requirements are enforced by smart contracts without any central authority.

1 Introduction

Over the past decade or so, the world of Health-Tech has become hugely popular and attractive for technologists and researchers. Widespread high-speed Internet availability and a plethora of smart devices available in the market today are the primary factors which has resulted in explosive growth in the sector. The ecosystem of the IoT is defined by many devices and sensors interconnected with each other via a communication and networking protocols. The capabilities and features offered by the ecosystem can be further enhanced by supporting technologies such as, Edge Computing, Cloud Computing, Data Analytics, etc. which are otherwise not possible by the limited computing and energy resources available within the devices itself.

By merging the two domains of Technology and Healthcare for end-users, it provides many benefits in the field of patient monitoring and delivering a quality service of healthcare. But with the benefits realized it also set of challenges related to Data Privacy and Security. Since the ecosystem is responsible for handling sensitive and personal data related to a user's health, it becomes imperative to protect this data from malicious use as any tampering of this data may even result in a patient's death.

Since most devices and sensors have limited resources in terms of compute and energy, it is important to deploy supporting technologies for providing the security and privacy-preserving capabilities to the ecosystem. Some researchers have proposed the use of Cloud Computing but it doesn't meet the low-latency requirements of the Health-Tech ecosystem. This can be resolved by introducing Edge Computing in the ecosystem but that again leads to other security concerns specially during data transmission over unsafe communication network as these are more prone to external threats and attacks.

1.1 Research Motivation

Blockchain or the Distributed Ledger is one such technology which has the potential to many of the security and privacy concerns faced in Healthcare Data Management. In contrast to other approaches, the Blockchain offers security and privacy protections without the need for powerful computing resources. A typical blockchain uses a distributed data storage and sharing system though an immutable and time-stamped public ledger system which is cryptographically secured. This ensures that any data stored on the blockchain network is immutable and irreversible which is very crucial for sensitive health data. Also, the decentralized nature of a blockchain network eliminates the need for employing the services of a third-party to manage the data storage requirements of the ecosystem.

The literature review brought forth gaps in the current solutions existing in the Healthcare Data Management through Technology. The issue of access control for private health information remains largely unresolved. This served as inspiration to develop and put forward a proposed architecture for healthcare data management based on the immutability and distributed properties of the permissioned blockchain network. The suggested architecture will put a strong emphasis on distributed, patient-centric, immutable security and privacy aspects, as well as data access control. The patient should have complete control over how, where, and by whom their data is accessed. I will also aim to address the common blockchain network's performance bottlenecks using the proposed architecture.

1.2 Research Question

How can the Healthcare Data Management leverage the Blockchain Network to implement Privacy Preserving Mechanisms for securing the sensitive and PII (Personally Identifiable Information)?

1.3 Report Structure

This report has been organized into following subsections. Section 2 covers the related work and Literature Review. Section 3 mainly focuses on the Research Methodology used in our work while Section 4 explains the Design Specifications of our proposed solution. Section 5 will focus on the Implementation of the proposed solution and Section 6 will focus on its Evaluation. Finally, Section 7 will cover the Conclusion and Future Work.

2 Literature Review

The focus of this section is to research and review the existing studies / research work conducted in the field of Data Security and Privacy for Technology in Health as well as various issues and challenges faced by the ecosystem. This review aims to analyse the covered aspects, proposed solutions, and limitations, if any, of the related works under consideration.

The section of Literature Review has been divided into three sub-sections, i.e., (i) Review of Technology in Health, (ii) Review of Data Security, (iii) Review of Data Privacy.

2.1 Review of Technology in Health

(Farahani et al., 2018) in their study, looked at how the healthcare industry has changed from being dependent on clinics to being patient-driven. The authors focussed on certain important health-related applications and systems, such as those that detect anomalies and provide early disease warnings. Data security, system scalability, privacy, system interoperability, and data standardization were among the issues they highlighted as obstacles on the road to ecosystem maturity.

IoT and cloud computing were coupled in a survey by (Darwish et al., 2019) for the healthcare data management. The writers explored the use of technology and Cloud Computing in smart health devices. They attempted to expand on the problems related to Technology in Health, such as data management, system scalability, storage, and system durability. Nevertheless, neither the study procedure itself was well planned nor the taxonomy of papers clearly defined.

A study by (Qi et al., 2017) was primarily designed to establish an IoT-based personalized healthcare ecosystem. The study outlined the application, network, sensor, and data management layers in depth. However, the writers failed to properly outline the specifics of the upcoming work and difficulties / problems with the suggested architecture.

In their research paper, (Ray et al., 2019) investigated the edge computing-based system solutions and protocols that were in use at the time. The authors suggested a multi-tiered architecture based on edge computing and compare the bandwidth and latency requirements for a cloud-based versus an edge computing-based architecture. However, the research article's coverage was not entirely comprehensive.

The Tech-Health ecosystem was further sub-categorized into sub-domains, such as Hospital Management, Out-Patient Systems, E-Health, and Mobile Health, in the study conducted by (Ahmadi et al., 2019). The lack of standardized designs and interoperability, according to the authors, is one of the Tech-Health industry's biggest problems. The authors' failure to include papers published between 2017 and 2020 in their work was the study's primary flaw.

In their study, (Kadhim et al., 2020) investigated the Health Monitoring Systems that uses the Internet as its communication backbone. The authors investigated the use of smart devices in the healthcare industry and how it could raise the standard of medical care by giving patients the right diagnosis.

2.2 Review of Data Security

A study on the problems with data integrity and authentication was released by (Atzori et al., 2010). The authors proposed adopting certain softwares for access and data management, with a focus on strengthening privacy protections. The paper falls short when it comes to describing the security issues that the networks and devices must deal with.

The study by (Miorandi et al., 2012) offers a wealth of fascinating information about the security issues the ecosystem was then currently dealing with. They largely concentrated on the important components of Confidentiality, Data Privacy, and Trust but ignored other factors, like data integrity, authentication, and data access control.

For the Tech-Health ecosystem, (Sicari et al., 2015) suggested an open architecture with features categorised into CAA (C-Confidentiality, A-Access Control, A-Authentication). The primary flaw in the suggested strategy was the taxonomy's vagueness, which led to an illogical classification scheme for future research endeavours.

With regards to the security vulnerabilities that may occur owing to interactions between the components of the IoT architecture itself, (Sfar et al., 2018) proposed a framework for IoT Security with a focus on Access Control, Data Privacy, and Authentication. However, the authors failed to address the external security threats.

In their paper, (Aksu et al., 2017) provided a methodology for Risk Assessment that accounts for both high and low risk levels based on quantification scores governed by pre-defined metrices and formulas for both internal and external threats. However, they paid less attention to how sensitive and Personally Identifiable Information (PII) should be stored securely.

Data Security in IoT has also been the subject of several additional research and surveys. The IoT Security has been the subject of successful analyses by (Farooq et al., 2015), (Leloglu, 2016), (Hossain et al., 2015), and (Ogonji et al., 2020). A recurring topic in the literature review is that the success of the deployment of an IoT ecosystem depends greatly on the End User's confidence in the system and its ability to govern and secure the sensitive data from any unwarranted incidents.

2.3 Review of Data Privacy

IoT devices are fundamentally designed to gather, transmit, and analyse user data that may qualify as sensitive Personally Identifiable Information (PII). The data is extremely important and needs to be secured from all threats and negative circumstances because of its inherent sensitivity. The user should also have control over the collection, storing, and processing of sensitive data and be able to make decisions in this regard.

(Perera et al., 2016)offered a privacy-by-design approach and guidelines for the designers of devices to meet the privacy concerns voiced by end users of these devices. These were lacking in a real-world implementation method for the suggested ecosystem architecture which might expose the IoT ecosystem to a wide range of security and privacy dangers.

The difficulties in developing data privacy protections for the IoT ecosystem are discussed in more detail (Rosner and Kenneally 2018). The authors suggest employing "offline" data storage would be a practical method for safeguarding sensitive data as a solution to this problem. However, the location of the storage and the security of the location are constrained by the scalability of devices in the ecosystem.

The protocols that are currently available for securing the communication networks in the IoT ecosystem have been thoroughly analysed by (Görmüş et al., 2018). The writers focused on the IoT networks' potential security concerns, weaknesses, and undesirable attacks. The security issues that were present at each tier of the ecosystem's ecosystem were successfully identified. However, the security protocols ignore the possibility of any internal threats and fails to include any cryptographic key management capabilities.

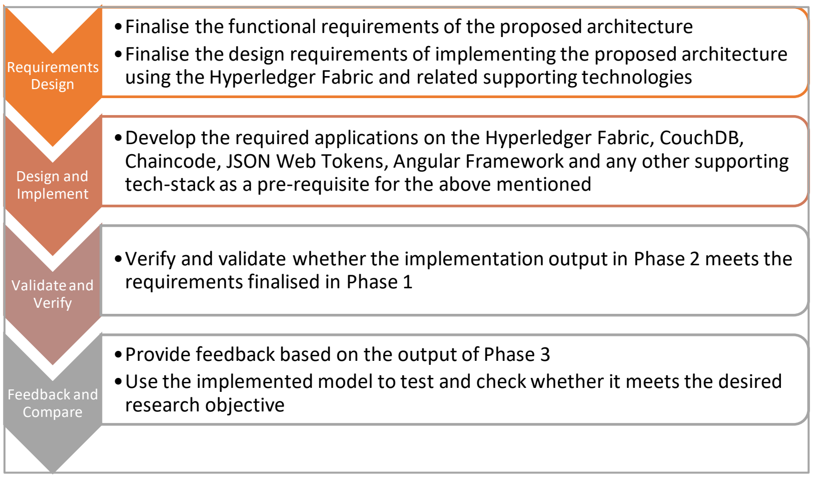
Research was done to determine how vulnerable certain IoT Health ecosystem devices are to both internal and external threats. The authors presented a framework based on the concept of secret cryptographic key sharing to address the issues related to data breaches and maintaining data integrity.

Most of the papers that have been evaluated above doesn’t address the issues related to Data Access and Privacy Management for sensitive user health data at its core architecture in the Tech-Health ecosystem. It has always been an afterthought which leaves the ecosystem open to more vulnerabilities as compared to inherent privacy preserving techniques. Our proposed architecture shall make use of core principles of Blockchain, viz., Distributed, and immutable to improve upon the security and privacy aspects of the Tech-Health ecosystem.

3 Research Methodology

In this section, we shall focus on the concepts and building blocks that shall form the basis of our proposed architecture.

The Research Methodology that is followed in this research work is based on the concept of Agile Software Development Process. Agile Methodology reduces obstacles and dependencies which may otherwise hinder the research process. It promotes quick learning and iterative build activities by utilizing continuous feedback mechanisms. It helped in identifying gaps in the research activity early during the process without impacting the timeline of the research project.



- Finalise the functional requirements of the proposed architecture
- Finalise the design requirements of implementing the proposed architecture using the Hyperledger Fabric and related supporting technologies

- Develop the required applications on the Hyperledger Fabric, CouchDB, Chaincode, JSON Web Tokens, Angular Framework and any other supporting tech-stack as a pre-requisite for the above mentioned

- Verify and validate whether the implementation output in Phase 2 meets the requirements finalised in Phase 1

- Provide feedback based on the output of Phase 3
- Use the implemented model to test and check whether it meets the desired research objective

Figure 1: Overview of Research Methodology

The research methodology also enabled us to deep dive into the evaluation of blockchain technology and its application in the healthcare sector via the Hyperledger Fabric. Prior to presenting the testing data in the form of scenarios that are implemented to test the defined needs, we first establish the major requirements of healthcare sector. Finally, we identified the specific technologies that were used to create the suggested blockchain environment on Hyperledger Fabric. It is important to establish that the goal of this research activity is to identify key requirements of implementing the blockchain / Hyperledger Fabric based architecture in the Healthcare sector and validate only those key requirements against a set of tests. A full production-ready application is not being developed as part of this research work.

3.1.1 Ethical Considerations

To maintain and adhere to the data ethics requirements, the research activity did not involve collection of any Personally Identifiable Information (PII) or any other kind of sensitive user-related data which could be compromised. The research activity uses dummy/mock data and fictional names of the entities involved only for the purpose of demonstration. The evaluation also does not involve human participants at any stage. This ensured that the research project is free of any potential ethical or data privacy issues.

4 Design

A blockchain based architecture offers many benefits as compared to a traditional centralised database storage system. The key benefits, out of the many, is the decentralized storage of data in multiple distributed computing nodes or peers which are cryptographically secured and encrypts the data. These nodes or peers are nothing but simple computer systems with specially designed softwares and computing environments. The same version of data is replicated on these peers which ensures the availability of data and eliminates single point of failure. The data cannot be deleted and only new versions can be created with the agreement of participating nodes which provides the benefits of immutability.

A more access-controlled version of blockchain, i.e., Hyperledger Fabric offers even more advantages for a sensitive scenario, such as, Healthcare Sector. The Hyperledger Fabric network offers a higher degree of permissioned and closed network unlike the open network (e.g., Bitcoin, Ethereum, etc.) and hence, no person can be added to the network without the prior agreement of the participating nodes based on a consensus mechanism. This provides confidentiality and access-control mechanisms to the sensitive user data. The smart contracts are known as Chaincode in a Hyperledger Fabric. The Chaincode contains the business logic which are packaged and deployed on to all participating nodes or peers in the network. The functioning of a Hyperledger Fabric network and its entities are explained in detail in the following sections.

4.1 Hyperledger Fabric – An Introduction

In the last couple of decades, Blockchain has gained quite a lot of prominence because of Bitcoin. It is a decentralized peer-to-peer network of distributed ledgers which consists of a chain of blocks. Each block contains the hash of its previous block which results in the formation of a chain. The typical composition of a block is header and a body which includes the block metadata. The metadata is usually made up of timestamp, hash of preceding block's header and nonce (which is cryptographic). A blockchain architecture is made up of following core components:

- Peer: This is also known as Nodes and is usually a device or a computing system within the blockchain network. It is used to store a duplicate version of all the transactions.
- Smart Contract: A set of business logic or transaction details
- Mechanism of Consensus: A set of rules and business logic which defines how a blockchain transaction is to be executed and agreed upon by the participating nodes.
- Transaction: It represents the act of information being transferred between the two blockchain addresses.
- Block: Used to store the batch of transactions which is distributed to all the participating peers in the network.
- Chain: It is the sequence of blocks ordered chronologically.
- Miners: Pre-designated peers or nodes which perform the operation of block verification

A typical blockchain network is permission-less, which means that anyone and everyone with an internet connection is allowed to participate and transact on the public blockchain network. As such, a public blockchain network never gained much popularity within enterprise application domain. This led to the development of permissioned and private blockchain network which puts limitations on who can transact on the network and as such is more suitable for enterprise needs or domains where data sensitivity is of utmost importance, i.e., Healthcare, Finance, etc. Hyperledger Fabric is one such example of a permissioned private blockchain network and forms the basis of our proposed architecture for sharing of healthcare data between participating entities.

The Hyperledger Fabric project was started by the Linux Foundation around the year 2015 with the aim of providing the businesses around the world with a network of distributed ledgers to execute sensitive as well as public transactions through the same network. It consists of a shared network and a trustworthy subnet inside that network that only permits participation from trusted parties. The various subnets within the umbrella network enables the users to communicate with other entities on the network with varying levels of confidentiality depending on the situation.

Hyperledger Fabric follows the concept of endorsement policies for approving a transaction by the pre-defined participating nodes. A Hyperledger Fabric architecture is made up of similar components as a typical blockchain network with some variations as outlined below:

- MSP: As the name suggests, it is the Membership Service Provider (MSP) and is responsible for enrolling and managing the identities and roles for entities on the network. It is responsible for authenticating any new client or node that wants to join the Hyperledger network.
- Endorsement Policy: It is the consensus mechanism, but unlike blockchain, only a set of pre-defined nodes or peers are responsible for approving a transaction before it is added to the ledgers.
- Chaincode: A smart contract is known as a Chaincode in the Hyperledger Fabric network which defines every business logic or rules on the network. These smart contracts can usually be written in any programming language, such as, Java, Go, etc.
- Channel: It is the permissioned network which is accessible only by the pre-approved participants and on which the transactions are executed.
- Nodes / Peers: There are two types of nodes or peers in a Hyperledger Fabric network. The peers which are responsible for approving a transaction (based on Endorsement Policy) are known as the Endorsing Peers which the peers which add the endorsed transactions to the distributed ledger are known as Committing Peers.
- Orderer: It’s a type of peer which acts as the central hub for communication on the Hyperledger network. It is responsible for the creation of blocks and maintaining a Ledger State which is consistent across the network.

4.1.1 Benefits in the Healthcare Sector

Hyperledger Fabric proves a worthy candidate to address the privacy and security issues for handling of sensitive user data in the Healthcare Sector. Since the blockchain network in Hyperledger Fabric is permissioned, it allows only pre-authorised trustworthy entities to execute and access the transactions on the network. Since the network is decentralised, it provides high availability as it eliminates any Single-Point-of-Failure (SPOF). This also enables easy exchange and transfer of information between entities which allows the patients to move to another health institute more conveniently.

The sensitivity of user's health data requires a higher degree of security and access control in the system which is enabled by the cryptographic capabilities of the Hyperledger Fabric. Chaincode or the smart contracts provides access to the distributed ledger which can be embedded with core business logic and functionalities of the targeted system. Some examples of this functionalities include, a doctor requesting access to patient's medical history, or a patient approving a data access request by the doctor, etc. A Hyperledger Fabric can be controlled by multiple chaincodes deployed on the network which provides modularity to the system by segregating different functionalities into different smart contracts. It also provides easy maintenance as a new chaincode can be deployed to the network whenever a new requirement needs to be implemented to the system.

The data collections are private in the Hyperledger Fabric wherein hash of the data is distributed across the participating nodes of the network. Whenever an organization or entity accesses the data, it can verify the integrity of the received data by comparing the value of stored hash with the received hash. This provides the privacy and security to the sensitive data through encryption.

The proposed architecture is also highly scalable as it is easy to configure the Hyperledger Fabric network with multiple entities. For example, when designing the system for healthcare, new entities can be added and configured for various roles in the ecosystem. There can be doctors, hospitals, medical officers, patients, health authorities, testing labs, and so on. The functionalities of each entity and access-level can be defined via Smart Contracts.

4.2 Architecture

The proposed architecture for the Healthcare Data Management consists of following components:

- Communication Channel (CC)
- Membership Service Provider (MSP)
- Certifying Authority for participating hospitals (CA)
- One peer/node per hospital
- Ledger per node (L)
- Chaincode or Smart Contracts deployed on each peer (SC)
- Orderer Node (ON)
- Consensus Mechanism or the Endorsement Policy
- Web Application
- Entities, such as, doctors, patients, administrator etc.

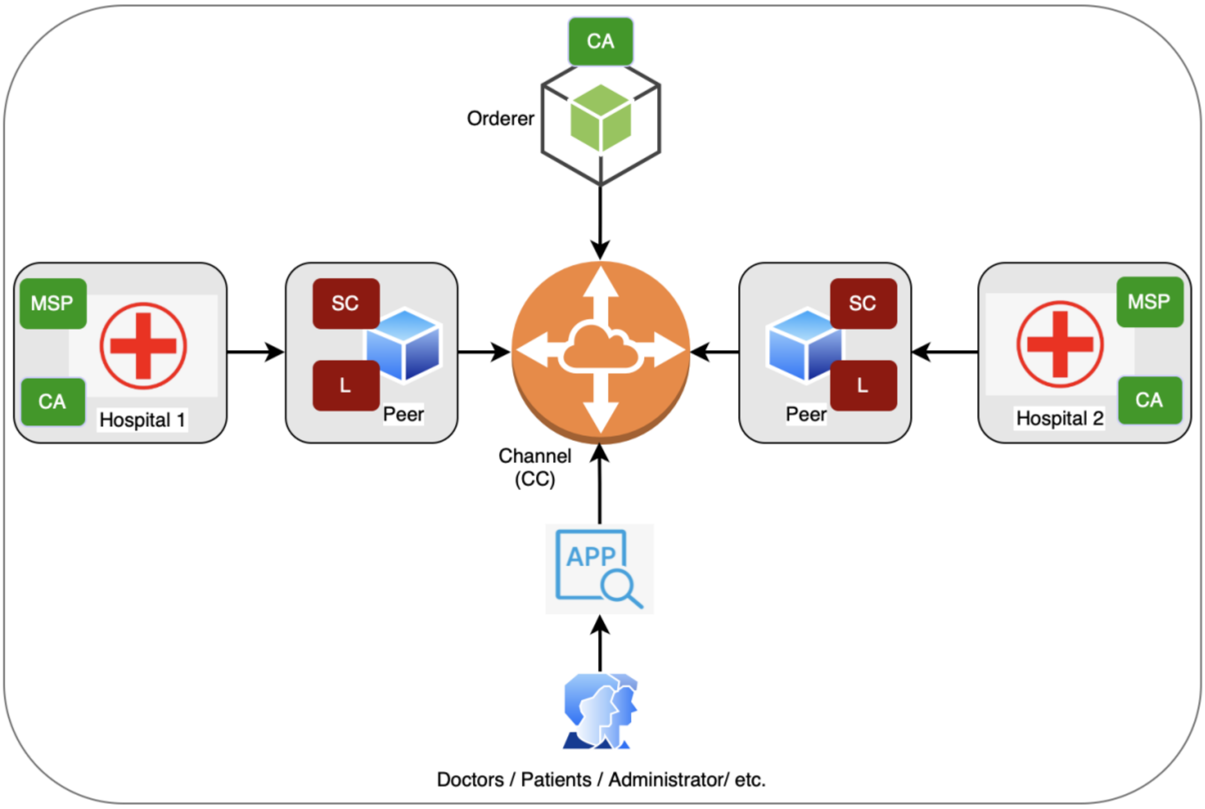


Figure 2 Proposed Architecture

For this research work and to keep it simple, only one channel has been included as part of the proposed architecture. This can be easily scaled up to multiple channels based on the requirements. All the entities and the participating peers/nodes will be using this channel to perform any transaction. Similarly, the architecture only depicts 2 hospitals with their own individual MSP and CA which provide the public-private key pairs for the organizations and sign the security certificates. The hospitals also have their peer nodes to interact with the channel, with each peer hosting a copy of the database. Any changes or modifications to the database, i.e., execution of a transaction, is broadcasted and updated on all the peers connected to the network. All kinds of interactions on the channel are dictated by the Smart Contracts, packaged, and deployed as Chaincode on all peer nodes in the Hyperledger Fabric network.

Whenever a peer validates a transaction, the ordering service decides what to do next based on the endorsing policy. Although we have implemented only one orderer node in our proposed architecture, a production-ready system can have multiple orderer nodes to increase the fault-tolerance. Our proposed architecture has assumed the following actors for demonstration purposes: a doctor, a patient (for whom the medical records are created) and an administrator who will be responsible for registering the patients and doctors on the system. These actors will interact with the Hyperledger Fabric network via a web application.

4.3 Privacy and Security Considerations

In the healthcare sector, ensuring the privacy and security of the user's sensitive health data is critical. To meet the security and privacy requirements, we have used Private Data Collection in our proposed architecture. The hospitals (organizations in the Hyperledger network) use the Private Data Collection mechanism to handle all the user's sensitive data on the distributed ledger. As explained earlier, the data is not visible to the participating hospitals but rather a hash value of the said data is shared with the hospitals. The data is stored privately only on the nodes of the hospitals which are allowed to access it. Whenever an unauthorized hospital tries to access that data collection, the Chaincode deployed on that node will verify the MSP-ID of the requesting hospital and then deny the access. The authorized MSP-IDs for this private data collection are defined in the collection configuration contained within the deployed chaincode package. These data collections must be defined for all the participating hospitals in the network. Since the configuration of these data collection is contained within the deployed Chaincode package, hence, the collections must be defined before the chaincodes are deployed on the Hyperledger Fabric network.

5 Implementation

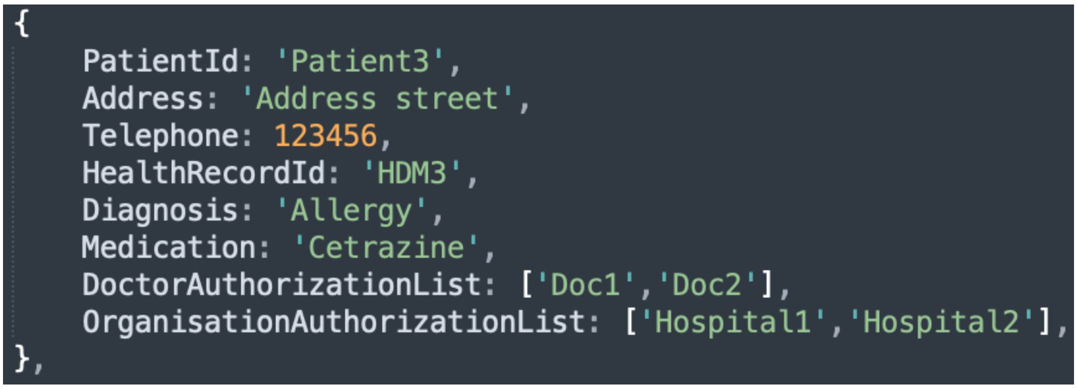
5.1 Hyperledger Fabric

Our proposed architecture for Healthcare Data Management is based upon the private and permissioned blockchain network, i.e., the Hyperledger Fabric network. The hospitals interact with the distributed ledger via the channel. For the sake of simplicity and demonstration, as explained earlier, we have considered only one channel in our proposed architecture which can be easily scaled up to multiple channels based on the requirements. In our test set-up, we used the default test-network provided by the Hyperledger Fabric. The docker images and accompanying certificate files were modified to use the default test organization as the hospitals. The configtx.yaml file and the accompanying Certificate Authority was modified in the docker-compose to include the hospitals. After these changes and their references are implemented in the respective files, the network is started which results in the creation of all peer nodes and their organizations.

5.2 Distributed Ledger

Based on the official repository and support website of the Hyperledger Fabric, it currently supports two different databases. The default database is the LevelDB wherein the data is stored as pairs of key and value. The second database supported by Hyperledger Fabric is the CouchDB which uses JSON to store the data. CouchDB is a more feature rich and popular database on the Hyperledger Fabric network as it offers richer queries to access data stored in JSON documents.

Hence, our proposed architecture uses CouchDB to store the patient's sensitive health data in a JSON document. Each patient is identified by their ID and acts as the owner of that record. A record might contain various fields such as patient's name, contact details, diagnosis, treatment, authorized doctors, etc. A sample data record in the JSON document is shown below:



```
{
    PatientId: 'Patient3',
    Address: 'Address street',
    Telephone: 123456,
    HealthRecordId: 'HDM3',
    Diagnosis: 'Allergy',
    Medication: 'Cetrazine',
    DoctorAuthorizationList: ['Doc1','Doc2'],
    OrganisationAuthorizationList: ['Hospital1','Hospital2'],
},
```

Figure 3 Sample Data Record

Each peer node has a running instance of CouchDB via a docker-image. Which means that the number of CouchDB images depend on the number of peers in the Hyperledger Fabric network.

5.3 Chaincode

As explained earlier, in the world of Hyperledger Fabric, smart contracts are termed as Chaincode. It contains all the business logic and rules as part of its code. Any and every transaction on the Hyperledger Fabric network is executed based on the business logic contained within the chaincodes. Typically, a smart contract can be written in any programming language. We have implemented one functionality per function in the smart contract which is written in Javascript for our proposed solution. Some of these functions are mentioned below:

1. A function to create a fresh health record whenever an administrator enrols a new patient on to the system
2. A function for updating a patient's personal information
3. A function for updating the health data of a patient
4. A doctor wanting to read a patient's medical record
5. A patient wanting to read their own medical record
6. Fetching a patient's past medical record
7. A patient can grant or revoke a doctor's access to their medical records. This is implemented by having an authorized doctor's ID list within each medical record in the JSON document. The patient can decide on granting or revoking access through the UI
8. A doctor can modify or view a patient's medical record through the UI if the doctor's ID is within the authorized IDs for that respective medical record

One chaincode maybe made up of multiple smart contracts deployed on the network. The steps to do the same are mentioned below and can be executed all at once using the deployCC command:

- Step 1: Chaincode is packaged
- Step 2: All peer nodes have the packaged chaincodes installed on them
- Step 3: The chaincode is approved for the participating hospitals or organizations
- Step 4: The installed chaincode package is committed to the communication channel

5.4 Hyperledger Fabric SDK

The Hyperledger Fabric software development kit has several components which can be configured to provide various features, such as, cryptography features, etc. There are various APIs provided by the SDK to perform various functionalities, such as, deploying smart contracts, executing transactions, etc. through UI. For our demonstration purpose, we have used the Node.js SDK provided by the Hyperledger Fabric network. All the functionalities mentioned in Section 5.3 are implemented using this SDK. These functionalities are triggered by various actors through the UI for the web application. The SDK is responsible for establishing and maintaining the connection with the test Hyperledger Fabric network via the provided Gateway. The SDK establishes the network connection by accessing the actor's and the channel's details from the wallet to direct them to the concerned peer and respective ledger. It also invokes the appropriate chaincode deployed on the network. Once the transaction is executed, the SDK terminates the connection using the respective methods in the same Gateway. Some of the core components of the SDK are mentioned below:

1. **Hyperledger Wallet:** It is used to store the metadata for the Hyperledger Fabric network, the public and private keys to be used by Certificate Authority. We have used file-type wallet for our implementation. The SDK uses MSP-ID in the wallet for authorizing the access level and connecting the user to the requested channel on the network.
2. **JSON Web-tokens:** These are used for authorization and session management of the user via APIs. The web-tokens are responsible for ensuring that the logged-in user is the same as the user who requested an API call. A JSON web-token is created when a user logs in using their ID and password. It is then encrypted and stored at the client. When a user requests for an API call, the SDK verifies the web-token and its integrity. It ensures that it is the same user who has logged in to the network. For demonstration and simplicity, we have stored the user credentials in a JSON file which, otherwise in a production-ready system, are encrypted and maintained in a different database for security reasons.

5.5 Web Application

Angular Framework has been used to develop the UI or the front end of the web application. We have created different views for different actors, i.e., administrator, doctor and patient based on their access. For Example, the View for registering or enrolling a new user on the system is available only to the administrators.

As explained in earlier sections, the server is responsible for creating a web-token once a user logs in. This token is then stored by the web application, in the browser, for user session management. Whenever a user makes an API call or tries to access the channel on the network, the server authenticates the hash of the web-token stored by the web application at the client browser. Accordingly, the server can authenticate the user, grant access to the network, and display the authorized views to the user.

Our web application consists of following views:

1. Login View: This view is common to all the actors and enables the user to enter their ID and password for accessing the system. If a user is already logged in, it shows the user to logout of the system.

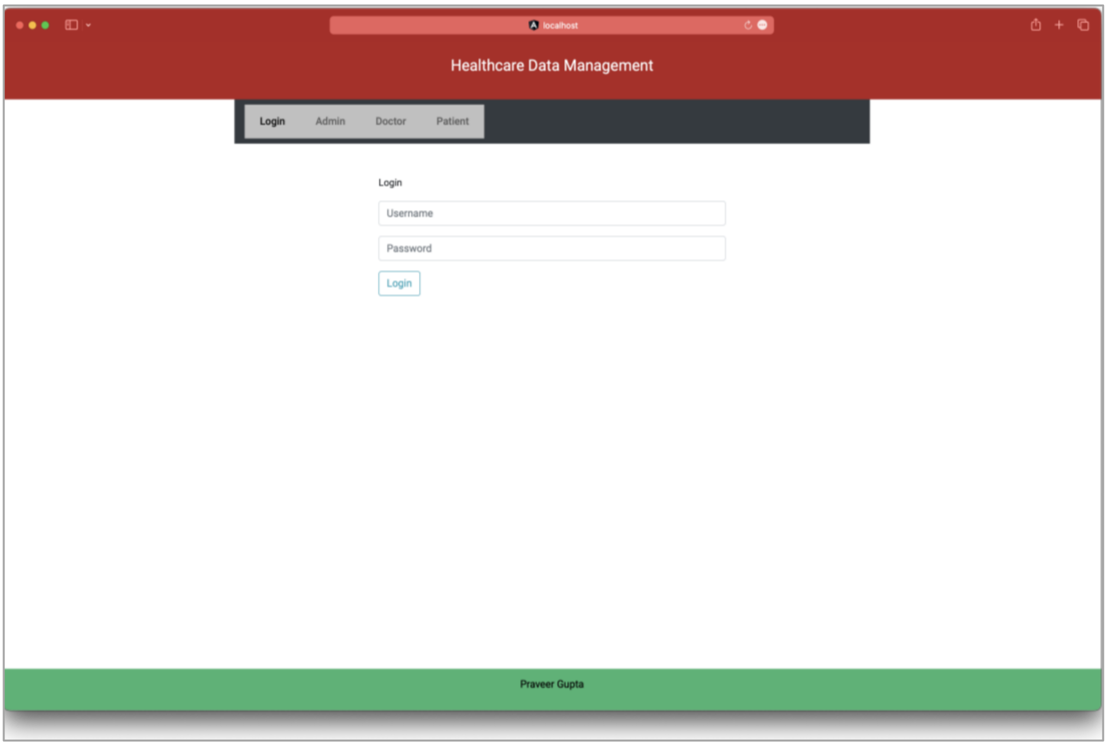
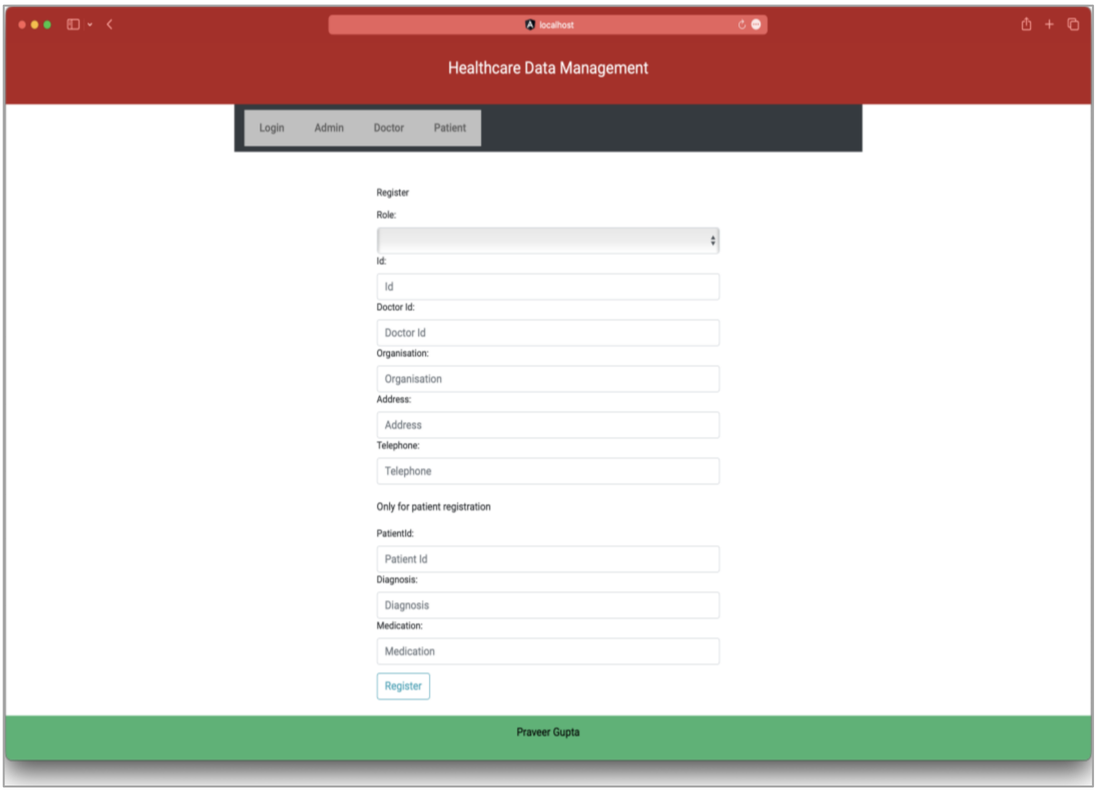


Figure 4 Login View

2. Administrator View: This view is exclusive for the administrator and is displayed only when the logged in user is identified as an administrator by their granted roles. It has the functionality to register a new user on the system. It could be further expanded to include new features, such as, modifying access levels, de-registering users, etc.



Healthcare Data Management

Register

Role:

Id:

Doctor Id:

Organisation:

Address:

Telephone:

Only for patient registration

PatientId:

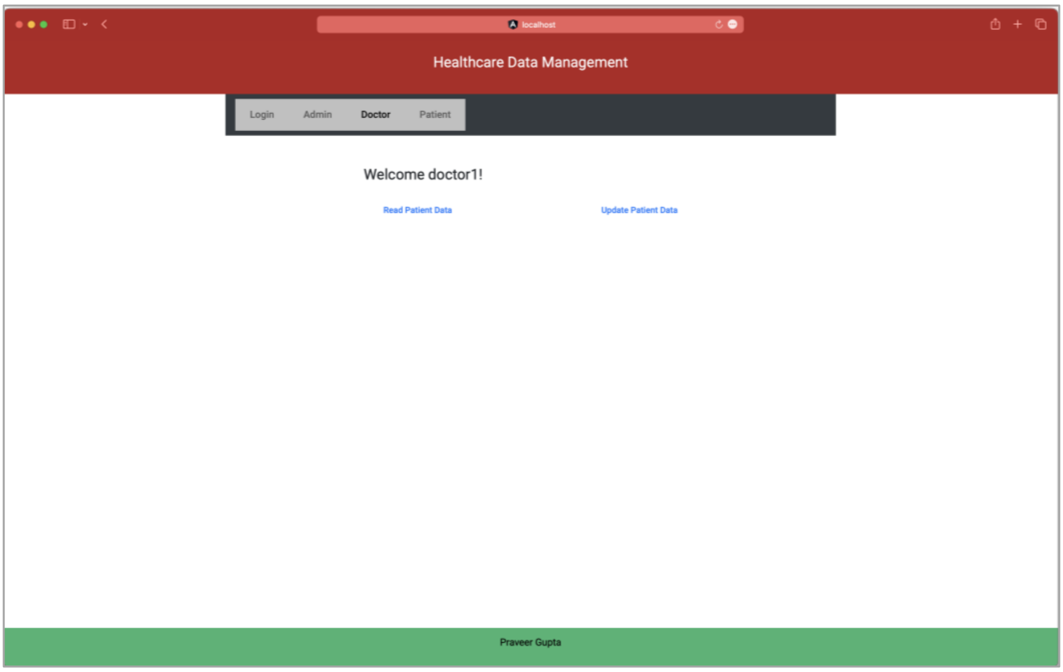
Diagnosis:

Medication:

Praveer Gupta

Figure 5 Admin View for Registering new Doctors / Patients

3. Doctor: This view is for the medical professionals or the doctors. It provides various features, such as, request patient's data, read patient's data and modify patient's data.

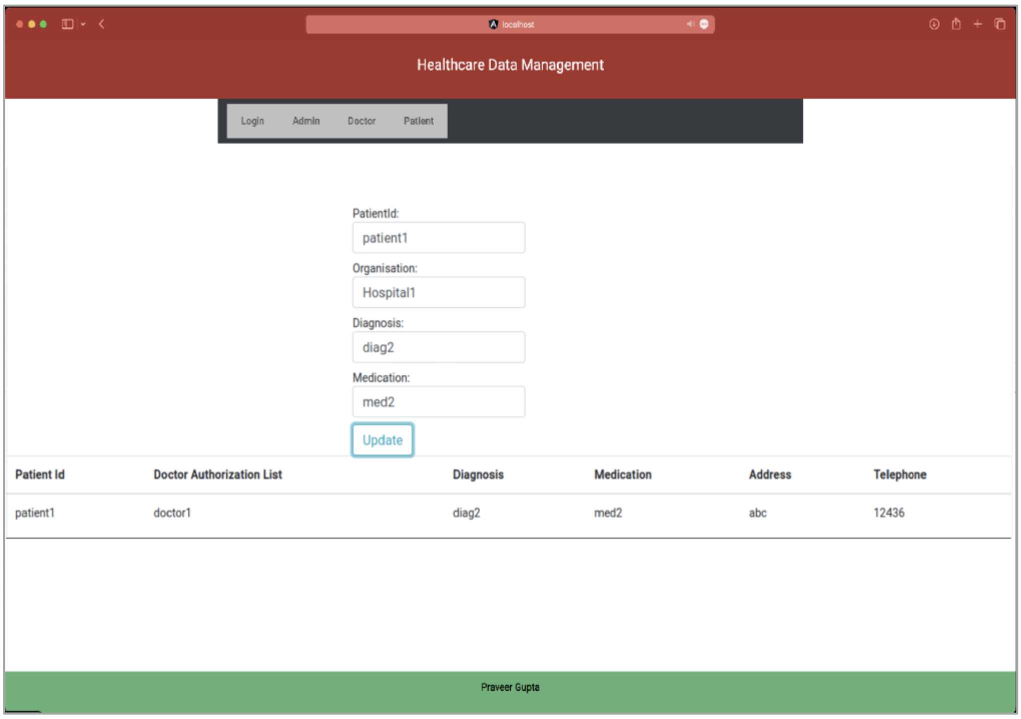


Healthcare Data Management

Welcome doctor1!

Praveer Gupta

Figure 6 Doctor's View



PatientId: patient1

Organisation: Hospital1

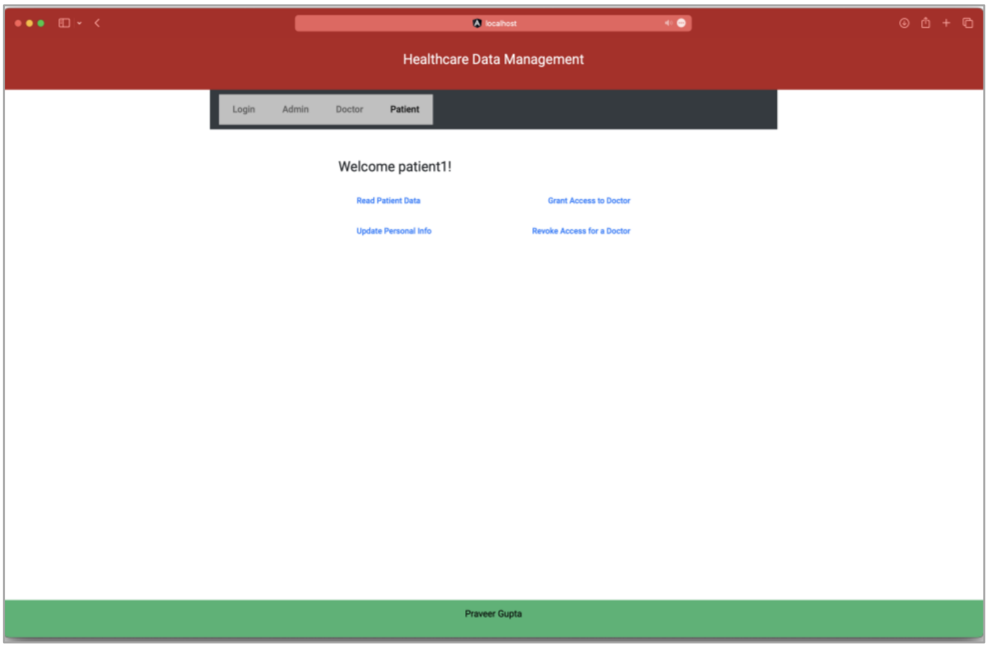
Diagnosis: diag2

Medication: med2

Patient Id	Doctor Authorization List	Diagnosis	Medication	Address	Telephone
patient1	doctor1	diag2	med2	abc	12436

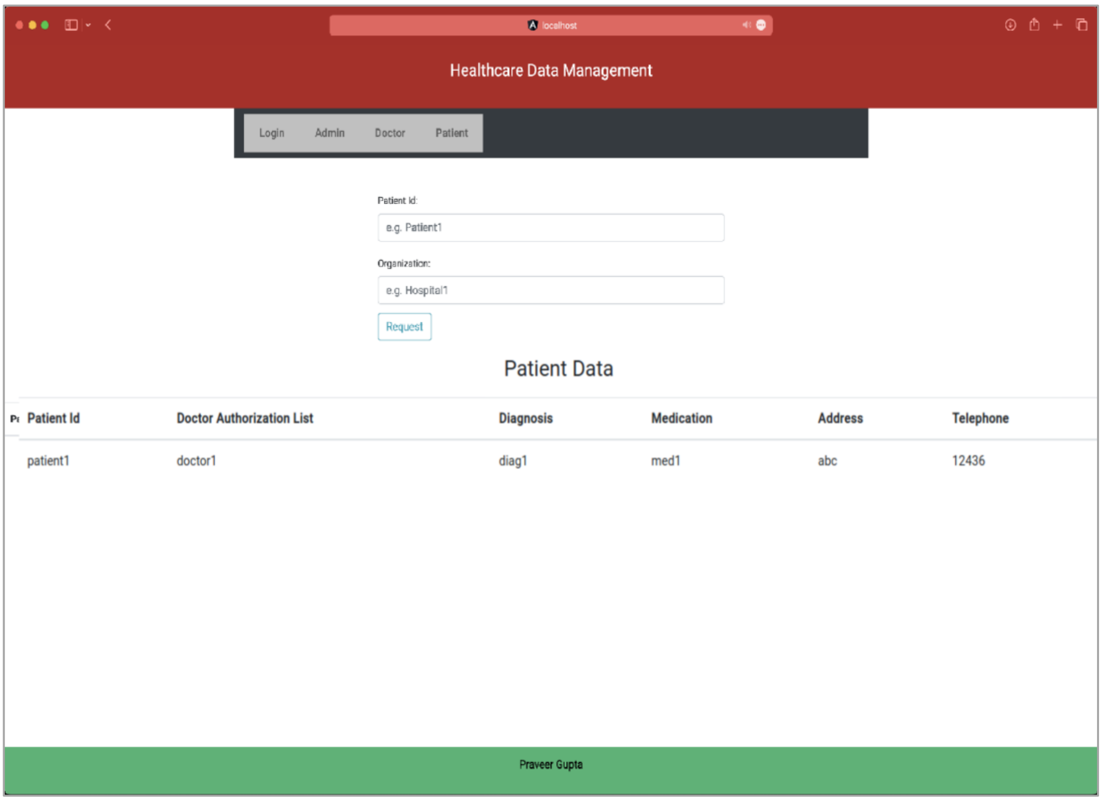
Figure 7 Doctor updating a Patient's Record

4. Patient: This view is for patients who can read their own data, update their personal information, grant, or revoke access to a medical professional or the doctor. The patient doesn’t have access to update their medical records.



Welcome patient1!

Figure 8 Patient's View



Healthcare Data Management

Patient Id: e.g. Patient1

Organization: e.g. Hospital1

Patient Data

Patient Id	Doctor Authorization List	Diagnosis	Medication	Address	Telephone
patient1	doctor1	diag1	med1	abc	12436

Praveer Gupta

Figure 9 Patient's Data view

6 Evaluation

6.1 Analysis

As part of this research activity, several scenarios were tested out, as explained in further sections, to thoroughly evaluate the proposed architecture. These scenarios have been broadly divided into following three categories:

6.1.1 Storage

This category deals with scenarios where doctors can create and modify medical records for registered patients. It also deals with patients being able to create and modify their personal records.

#	Scenario	Result
1.	An authorised doctor can create medical record for patient	Successful
2.	An authorised doctor can modify medical record for patient	Successful
3.	Patient can create personal information for themselves	Successful
4.	Patient can modify personal information for themselves	Successful
5.	An actor (doctor/patient) can be uniquely identified	Successful
6.	An encrypted medical record can be successfully viewed by authorised doctor	Successful

6.1.2 Data Privacy

It is extremely important for our proposed architecture to meet the security and privacy considerations required by a system which deals in a user's sensitive health data. Accordingly following test cases were created and tested to verify whether the proposed architecture meets these requirements.

#	Scenario	Result
1.	The homepage view depends on the user's role (administrator / doctor / patient)	Successful
2.	The doctor cannot view patient's data without patient granting the access	Successful
3.	The patient can grant access	Successful
4.	The patient can revoke access	Successful

6.1.3 Data Security

Data Security is another critical aspect which we have aimed to address in our proposed architecture. To have enhanced security levels in the system, we have used Private Data Collection (PDC) feature provided by the Hyperledger Fabric. Through Private Data Collection (PDC), a Hyperledger Fabric network maintains the privacy of data between different organizations on the network. PDC enables the organizations to see the transactions being executed on the network but only the hash value of data set can be viewed rather than the data itself. This is achieved by the fact that the data is stored in a database that is private to the organization which has the permission to access it. The smart contract (deployed on the network) will check the mspID and other information and will refuse access if any other user or organization that does not already have access to that collection tries to access it. The Private Data Collection (PDC) uses the collections_config file written in JSON. When deployed with chaincode, this file has the details about configuration for creating PDCs. It also contains the details about the endorsement policy.

6.2 Review of Research Objectives

The objective of the Research Activity was to propose an architecture for Healthcare Data Management that has Data Privacy at its core rather than an afterthought. As a result, we proposed an architecture and developed a system (only for demonstration purposes) based on the Hyperledger Fabric, which is a permissioned blockchain network. Information sharing is secure, and interoperability makes it easier for organizations to implement this system. The system's transparency in data handling and security benefits both the medical personnel and the patients.

However, compared to a more conventional approach, the costs of implementing the proposed architecture in a production real-world environment would be higher. Since this system is based on a Permissioned Blockchain Network, systems would need to be bought and installed at specific locations. Personnel would need to be hired who have the technical knowledge and skillset to operate these systems. But over the course of time, the cost of adding new sites / organizations / hospitals would be less as the initial infrastructure would have already been set up.

The proposed architecture also guarantees the three fundamental elements of information security: availability, integrity, and confidentiality.

- Using a private and access-restricted blockchain network, such as, Hyperledger Fabric ensures **Data Confidentiality**.
- The Hyperledger Fabric achieves **Data Integrity** through Hashing and using a distributed database, i.e., CouchDB.
- **Availability** is achieved by the inherent fault-tolerant properties of the Hyperledger Fabric network as it reduces the count of failed connection to the data nodes. Since, it uses a shared and distributed database, i.e., CouchDB, it maintains multiple copies of the data at different participating nodes on the network.

7 Conclusion and Future Work

7.1 Conclusion

Proving to be one of the most appealing technologies of the recent times, Hyperledger Fabric offers many advantages for sensitive Healthcare Data Management. Through features, such as, Private Data Collection, Smart Contracts, etc., it provides a more secure and private data management environment. In the Healthcare Data Management, it enables multiple hospitals, medical professionals, and patients to co-exist in the same system without infringing on any stakeholder's privacy. The data sharing between the authorised parties is easier and more secure with the control of the data residing with the owner of the data i.e., the patient. They are not required to carry the physical copies of their medical records with them even if they move or switch to a different hospital/doctor. This gives a boost to the digital revolution in healthcare and makes everyone's life easier without compromising on privacy or security.

7.2 Future Work

As is always the case with any technology, improvements and additions can be made to the proposed architecture to make it more suitable for the real-world use cases. Some of these improvements are listed down below:

1. The architecture can be scaled up to include more organizations, i.e., hospitals. This would mean more peers / nodes connected to the Hyperledger network resulting in higher transaction volumes. Hence, more ordering peers and communication channels will be needed to manage and handle the requests load. This could be achieved by using an open-source platform, such as, Apache Kafka to manage event streaming in a distributed environment.

2. The architecture can be modified to include metadata concept on the network. As the system scales up, the volume of data stored on the blockchain will increase exponentially which may lead to network slowdown. Hence, the distributed ledgers on the blockchain may be targeted to store only the metadata while the actual data is being stored on traditional database systems. In this way, the architecture can make use of the benefits offered by the permissioned blockchain network as well as conventional database systems without compromising on the performance.
3. The architecture and the web application can be upgraded to support reports and multimedia health records, such as, MRI scans, X-Rays, Blood Reports, etc.
4. The architecture can be integrated with sensors and smart devices, such as, smart watches, fitness trackers, etc. to provide real-time updates and patient monitoring.

References

Ahmadi, H., Arji, G., Shahmoradi, L., Safdari, R., Nilashi, M. & Alizadeh, M. (2019) The application of internet of things in healthcare: a systematic lite rature review and classification. *Universal Access in the Information Society,* 18(4)**,** 837-869.

Aksu, M. U., Dilek, M. H., Tatlı, E. İ., Bicakci, K., Dirik, H. I., Demirezen, M. U. & Aykır, T. (2017) A quantitative CVSS-based cyber security risk assessment methodology f or IT systems. IEEE, 1-8.

Atzori, L., Iera, A. & Morabito, G. (2010) The internet of things: A survey. *Computer networks,* 54(15)**,** 2787-2805.

Darwish, A., Hassanien, A. E., Elhoseny, M., Sangaiah, A. K. & Muhammad, K. (2019) The impact of the hybrid platform of internet of things and cloud comp uting on healthcare systems: opportunities, challenges, and open probl ems. *Journal of Ambient Intelligence and Humanized Computing,* 10(10)**,** 4151-4166.

Farahani, B., Firouzi, F., Chang, V., Badaroglu, M., Constant, N. & Mankodiya, K. (2018) Towards fog-driven IoT eHealth: Promises and challenges of IoT in medi cine and healthcare. *Future Generation Computer Systems,* 78**,** 659-676.

Farooq, M. u., Waseem, M., Mazhar, S., Khairi, A. & Kamal, T. (2015) Article: A Review on Internet of Things (IoT). *International Journal of Computer Applications,* 113(1)**,** 1-7.

Görmüş, S., Aydın, H. & Ulutaş, G. (2018) Security for the internet of things: a survey of existing mechanisms, protocols and open research issues. *Journal of the Faculty of Engineering and Architecture of Gazi Univers ity,* 33(4)**,** 1247-1272.

Hossain, M. M., Fotouhi, M. & Hasan, R. (2015) Towards an analysis of security issues, challenges, and open problems in the internet of things. IEEE, 21-28.

Kadhim, K. T., Alsahlany, A. M., Wadi, S. M. & Kadhum, H. T. (2020) An overview of patient's health status monitoring system based on Inte rnet of Things (IoT). *Wireless Personal Communications,* 114(3), 2235-2262.

Leloglu, E. (2016) A review of security concerns in Internet of Things. *Journal of Computer and Communications,* 5(1), 121-136.

Miorandi, D., Sicari, S., De Pellegrini, F. & Chlamtac, I. (2012) Internet of things: Vision, applications and research challenges. *Ad hoc networks,* 10(7), 1497-1516.

Ogonji, M. M., Okeyo, G. & Wafula, J. M. (2020) A survey on privacy and security of Internet of Things. *Computer Science Review,* 38, 100312.

Perera, C., McCormick, C., Bandara, A. K., Price, B. A. & Nuseibeh, B. (2016) Privacy-by-design framework for assessing internet of things applicati ons and platforms. 83-92.

Qi, J., Yang, P., Min, G., Amft, O., Dong, F. & Xu, L. (2017) Advanced internet of things for personalised healthcare systems: A sur vey. *Pervasive and Mobile Computing,* 41, 132-149.

Ray, P. P., Dash, D. & De, D. (2019) Edge computing for Internet of Things: A survey, e-healthcare case stu dy and future direction. *Journal of Network and Computer Applications,* 140, 1-22.

Sfar, A. R., Natalizio, E., Challal, Y. & Chtourou, Z. (2018) A roadmap for security challenges in the Internet of Things. *Digital Communications and Networks,* 4(2), 118-137.

Sicari, S., Rizzardi, A., Grieco, L. A. & Coen-Porisini, A. (2015) Security, privacy and trust in Internet of Things: The road ahead. *Computer networks,* 76, 146-164.