Elizabethtown College

JayScholar

Summer 2020

Building a Secure Inter-institutional Data Sharing Platform with Blockchain.

Grace Cuff

Jeff Edmonds

Building a Secure Inter-institutional Data Sharing Platform with Blockchain

Grace Cuff and Jeff Edmonds
Advisor: Dr. Jingwen Wang

Department of Computer Science
Elizabethtown College, Elizabethtown, PA
Email: cuffg@etown.edu and edmondsj@etown.edu

Abstract— Current cloud-based solutions to inter-institutional data sharing are expensive regarding technical resources like network bandwidth, often creating latency in real world scenarios. These network costs, as well as data storage costs, are often out of reach for most research budgets. Because the goal is to share inter-institutional data, the solution should use a consensus protocol to ensure that the data isn't manipulated or mismanaged in any other way. The solution should also include an easy to navigate user interface that is available through any web browser. The goal of this research project is to determine if the Hyperledger Fabric blockchain technology can be integrated with a user-friendly website to solve these problems concurrently.

Keywords— blockchain, data sharing, access control

I. INTRODUCTION

The problem this platform is attempting to solve is how can we make peer to peer data sharing more accessible and secure on a broader scale that is less resource intensive? Currently, patients are able to download their own data through patient portals offered by their doctors office and transfer data between doctors if necessary. However, there is no easily accessible single platform to securely share data with researchers and medical institutions. For example, Wellspan has a patient portal where

Because of the HIPPA Privacy Rule which "describes the ways in which covered entities like University Hospitals (UH) can use and disclose protected health information (PHI) for research purposes. Under the Privacy Rule covered entities may use and disclose PHI for research purposes with individual authorization"(University Hospitals). University Hospitals organization provides a way to submit your research credentialing application which, if approved, grants you access to University Hospitals PHI to be used for Institutional Review Board approved research studies (University Hospitals). The motivation for our data sharing platform was to analyze if these three aspects would: create a more efficient method of sharing medical data, benefit using blockchain-based framework rather than cloud-based, and provide a clean interface for requesting and authenticating data

II. DESIGN

The design has two main components, the website and the blockchain network (Figure 1). The website serves as a user interface for all participants in the network, including patients, administrators, and researchers. Behind the website are the server side applications, which are responsible for the blockchain functionality.
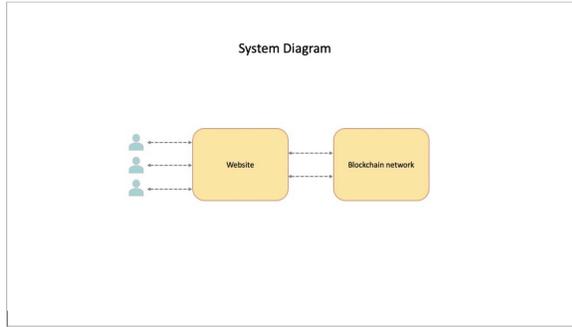
Figure 1. System Diagram

Figure 2. Basic Hyperledger Fabric Blockchain Network

Hyperledger Fabric was chosen for the blockchain technology, because it is permissioned, which allows only certain identities on the network to make changes. It is also highly customizable. Many blockchain technologies are made exclusively for financial data, but Fabric can be adjusted to suit other types of data [1].

A basic Hyperledger Fabric blockchain network consists of member organizations, where each organization is represented on the network by peer nodes. The peer nodes contain the important information for the network, including the ledger that contains the network data, and the chaincode which holds the computer code that processes data on the ledger. To access the network, users register through their respective organizations and hold their access IDs in a network wallet (Figure 2).

As of now, the chaincode applications are able to:

1. Instantiate a ledger of patient data.
2. Query the entire ledger.
3. Query a single patient entry in the ledger.
4. Add an entry to the ledger.

III.   USE CASE

In example (Figure 3), let's say a new patient will be using the system. They would start by creating an account via the website's user interface. The patient would sign in and give their permission for the blockchain network to access their data. When chan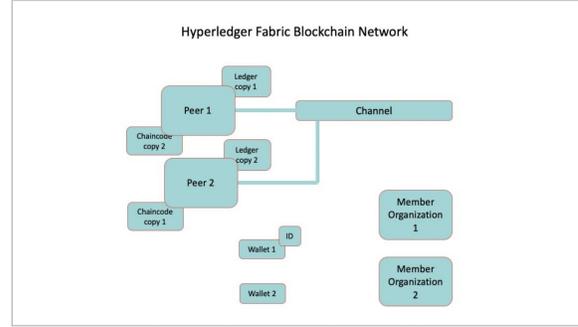ging data on the blockchain ledger, the blockchain network needs to use the chaincode applications located on the organization's peer node. For instance, if the patient wanted to add their record to the ledger, they would input it using the website. The website would call on the RESTful APIs as an intermediary that then calls on the chaincode application to add the new data.

Before adding the data, the chaincode would first need to propose a change to the ledger to another peer on the network. The other peer would endorse the transaction, both peers would update their ledgers with the new data. Lastly, the patient would be notified through the website.

IV.   BLOCKCHAIN REQUIREMENTS

A.   Blockchain Hardware Requirements

Testing of the system was done using a virtual host local network. This requires a Linux PC to host the local network and applications. For deployment, the blockchain network would require a Linux server with internet access.

B.   Blockchain Software Requirements

There are a number of software requirements for running the blockchain network server:

1. Ubuntu 18.04
2. Go 1.13.x
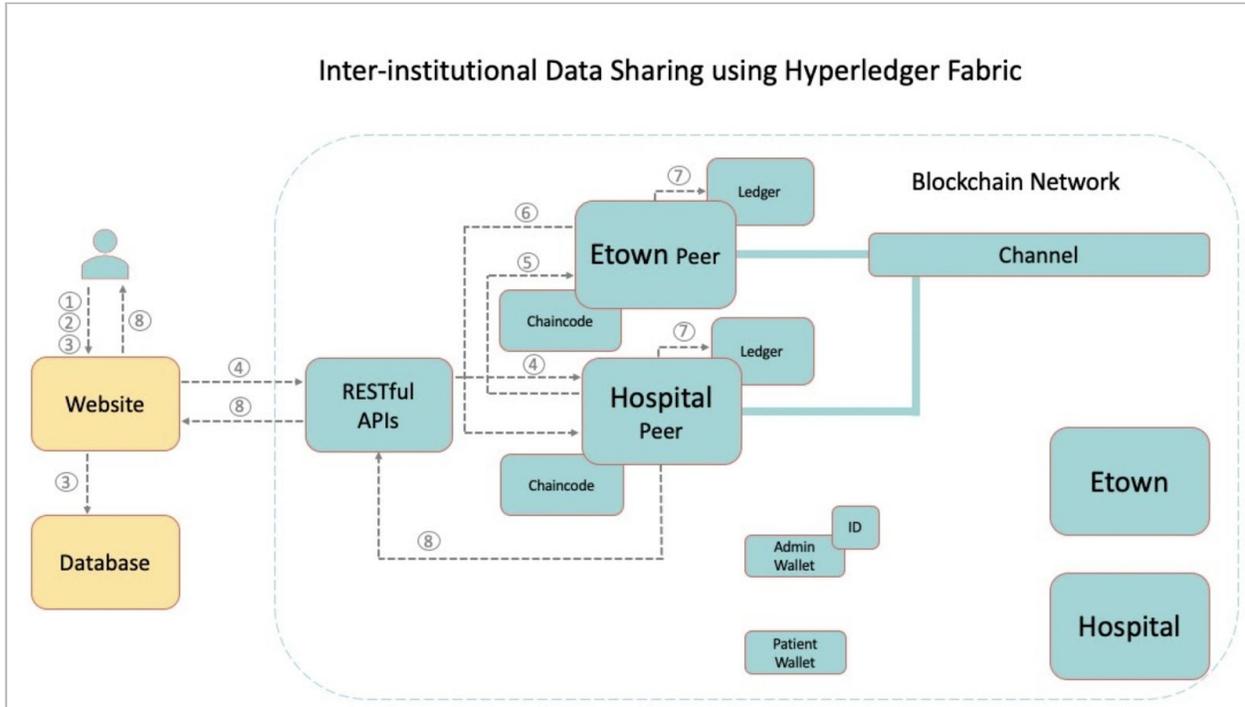3. Git client
4. cURL
5. Docker

Figure 3. 1. Patient signs in. 2. Patient gives permission. 3. Data added via website. 4. Website calls RESTful APIs. 5. Chaincode proposes change. 6. Transaction endorsed. 7. Ledgers are updated. 8. Patient is notified.

6. Docker-compose
7. Hyperledger Fabric 2.0.1 samples repository
8. Hyperledger Fabric-CA 1.4.6
9. Node.js
10. Node Package Manager

Ubuntu 18.04 Desktop is used for the operating system because it is highly customizable and offers long term support. The Desktop version of the operating system was used to facilitate the development process by providing a GUI. Many components of the Hyperledger Fabric samples repository are written in Go. Node.js is required because Javascript was chosen as the development language of choice and Hyperledger Fabric offers a Javascript SDK. Node Package Manager, git, and cURL are all required to download and install necessary Fabric repositories, Node packages, platform specific binaries, and so forth, which the Fabric system requires. NPM is also required for the Express and Body Saver packages that are used for the RESTful APIs [2].

V. FRONT END INTERFACE

### A. Software Requirements

The interface was coded using Python Version 3.7 and Django Version 3 in PyCharm. The website is run through the command line on the local host.

Django is a web framework designed to make website development easier and more secure. Because of the data that is being uploaded and shared on this platform, we need the security that Django offers.

B. Databases

Within the website, there are four main databases that store the following information

Databases
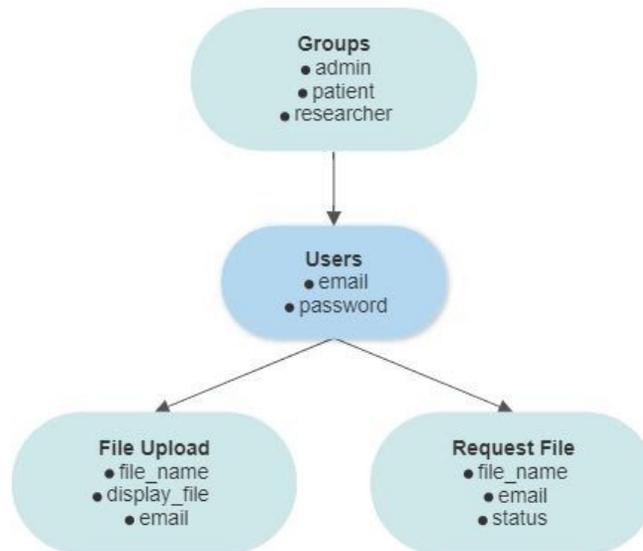
VI. FEATURES OF THE WEBSITE

A. Login/Authentication

Django provides us a secure authentication mechanism. Users are identified with their email and a password that is encrypted by Django's framework.

B. Homepage Display

After users login, their homepage displays a table with COVID-19 studies, users are able to click on the study title and view more information about them.

C. Uploading Files

Users are able to upload any files with a .txt or .xml extension, those files are then displayed in the data center, along with every file that has been uploaded to the website by other users.

D. Data Center

After users upload a file, it is stored in the data center. Users are only able to view the contents of their own files. Every file is displayed in a table, but a hyperlink is displayed only if the current email logged in matches the file owner's email.

E. Requesting Access

One of the most necessary features of this website is the ability to request access to a file. If a user does not have access to a file they can click "request access" and that will send a request to the user. The owner of the file can either accept or deny the request. If the request is accepted, the file will now include a hyperlink to view the file in the data center of the user who requested.

VII. CONCLUSION

By developing a frontend user interface and backend blockchain network server, we have shown that blockchain technology can be used with patients' electronic medical record data.

Once the back and front ends of the system have been fully integrated, the system will need to be tested in a real-world scenario to ensure that it can be scaled for efficient transfer of medical data inter-institutionally.

REFERENCES

1. Hyperledger Fabric, *A Blockchain Platform for the Enterprise*, Hyperledger Fabric, 2020. Accessed on:

Aug. 14, 2020. [Online]. Available:
https://hyperledger-fabric.readthedocs.io/en/release-2.0/

2. KC Tam. *An Implementation of API Server for Hyperledger Fabric Network*, Medium, Sep. 4, 2019. Accessed on: Aug. 14, 2020. [Online]. Available: https://medium.com/@kctheservant/an-implementation-of-api-server-for-hyperledger-fabric-network-8764c79f1a87

3. "Access & Use of Patient Records for Research Purposes." *University Hospitals. The Science of Health. The Art of Compassion.*, www.uhhospitals.org/uh-research/for-researchers/research-and-clinical-trials/for-researchers/frequently-asked-questions/access-and-use-of-patient-records-for-research-purposes.