A Practical Training Report on

HEALTHCARE AND SUPPLY CHAIN MANAGEMENT

Submitted in fulfillment of the requirement of the degree of

Bachelor of Technology
Degree in
COMPUTER SCIENCE ENGINEERING
By

Aditya Prakash Joshi
CSE/15/002

Department of Computer Science Engineering

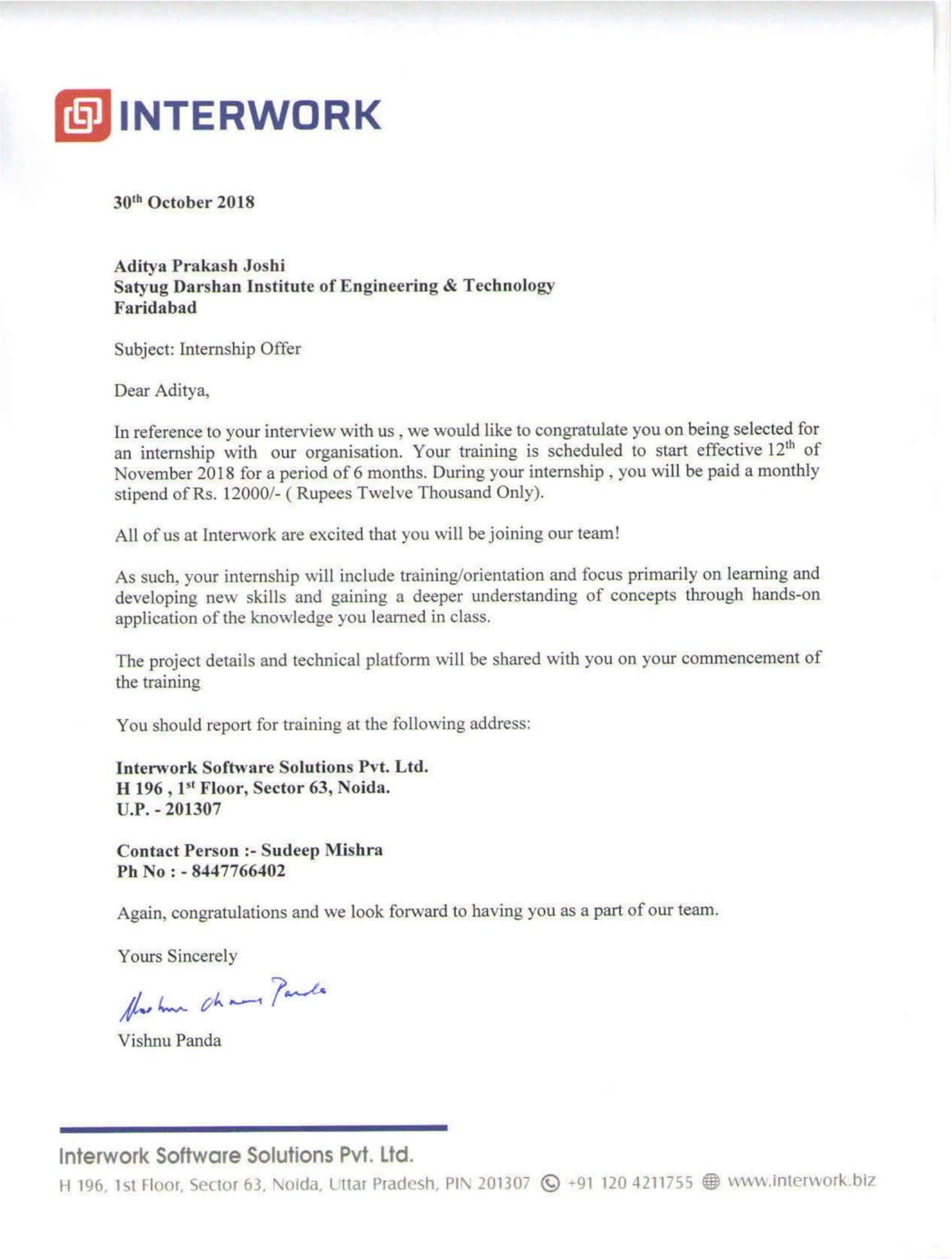
Satyug Darshan Institute of Engineering & Technology

Bhopani, Faridabad, Haryana, India

MAHARISHI DAYANAND UNIVERSITY, ROHTAK
MAY, 2019

1.TABLE OF CONTENTS

2. TRAINING LETTER OF COMPANY



INTERWORK

30th October 2018

Aditya Prakash Joshi
Satyug Darshan Institute of Engineering & Technology
Faridabad

Subject: Internship Offer

Dear Aditya,

In reference to your interview with us , we would like to congratulate you on being selected for an internship with our organisation. Your training is scheduled to start effective 12th of November 2018 for a period of 6 months. During your internship , you will be paid a monthly stipend of Rs. 12000/- (Rupees Twelve Thousand Only).

All of us at Interwork are excited that you will be joining our team!

As such, your internship will include training/orientation and focus primarily on learning and developing new skills and gaining a deeper understanding of concepts through hands-on application of the knowledge you learned in class.

The project details and technical platform will be shared with you on your commencement of the training

You should report for training at the following address:

Interwork Software Solutions Pvt. Ltd.
H 196 , 1st Floor, Sector 63, Noida.
U.P. - 201307

Contact Person :- Sudeep Mishra
Ph No : - 8447766402

Again, congratulations and we look forward to having you as a part of our team.

Yours Sincerely

Vishnu Charan Panda

Vishnu Panda

Interwork Software Solutions Pvt. Ltd.
H 196, 1st Floor, Sector 63, Noida, Uttar Pradesh, PIN 201307 ☎ +91 120 4211755 🌐 www.interwork.biz

3. INTRODUCTION TO PROJECT

HEALTHCARE AND SUPPLY CHAIN MANAGEMENT

Managing and securing data within healthcare and supply chain management are two great examples of principal concepts influencing and being impacted by possible blockchain adoption. Let’s take a brief look at each one:

Healthcare: Better data sharing between healthcare providers means a higher probability of accurate diagnoses, more effective treatments, and the overall increased ability of healthcare organizations to deliver cost-effective care. Blockchain technology can allow various stakeholders in the healthcare value-chain to share access to their networks without compromising data security and integrity, by allowing them to track data provenance as well as any changes made.
Supply Chain Management: One of the most universally applicable aspects of blockchain technology is that it enables more secure and transparent monitoring of transactions. With blockchain, the transactions can be documented in a permanent decentralized record — reducing time delays, added costs and human errors.

Innovating traditional models of healthcare

Blockchain has the potential to propel innovation in preventative care and community-based healthcare models. The capacity of a distributed ledger technology for ensuring data integrity while sharing between parties can ensure collaboration between rising trends in healthcare, which are vital to the improvement of health in communities worldwide.

Blockchain can tie together a complex team-based healthcare, finance and payment with the care provided along with it. The inherent properties of cryptographic public and private key access, proof of work and distributed data, creates a new level of integrity for healthcare information. Blockchain technology also makes it easy to track a drug as it moves from the manufacturer to the patient. This improves the traceability of a drug as it moves across the supply chain, and helps prevent drug counterfeiting.

Blockchain provides frictionless connectivity, strengthened by smart contracts and authorization to access all electronic health data. Its transaction layer can enable instantaneous access to a diverse set of standardized, anonymous and non-patient identifiable information. Transparency and automation can also lead to higher efficiency and lower administration costs. It is a phased approach rather than an instant overhaul of systems, and hence is suited to healthcare sector.

Educating the healthcare industry on security

In the current system, security and trust are the most common concerns shared by businesses regarding the information shared between different entities. Information can be entered anywhere along the line of communication and this leads to trust issues, especially in the healthcare industry. There are also concerns where multiple vendors hold different versions of the same patient record that are not validated, resulting in various errors, inconsistency and incompleteness. Add to that reports of security breaches, tampering of personal data and the ever-present hacking threat, it’s not surprising healthcare officials are concerned.
Since blockchains are cryptographically secure and the data present therein can be authenticated using digital signature that are unique to each person, this technology could be the answer to most of these concerns.

Scaling blockchain for healthcare

The healthcare industry is on the verge of disruption in its digital infrastructure. The current system does not fully support the security or interoperability that is inherently necessary. To utilize collected medical data to its maximum potential, data portability and interoperability of records between systems is a must.
With the advent of wearables and numerous new IoT devices that are interconnected with their data flows harnessed, better security is needed and readily accessible to healthcare professionals. All these challenges could be minimized with the help of blockchain technology and its interoperability, integrity and security, and portable user-owned data.

Advantages of Blockchain for Health Data

The blockchain era has already begun. Taking into account the fast progress in the development of new and more efficient healthcare record systems, wearable devices, and medical examination systems implementing artificial intelligence, cryptography will become an important part of the way hospitals work. There are, however, a few improvements still needed in order for seamless blockchain adoption across the entire medical industry. According to Hyperledger's survey, 42.9% of healthcare organizations suppose that the interoperability of electronic health records will help for faster blockchain implementation; with 28.6% of respondents ready to use this technology in care settings today. So, what are the benefits of blockchain technology in healthcare?

Data Provenance and Integrity

With an ongoing increase in patient numbers, healthcare providers have to manage more and more health data on a regular basis. As the data volume increases each year, it becomes harder for hospitals and clinics to process and store information.

Data managed by medical organizations includes:

· Patient health information (PHI);

· Electronic health records;

· Data collected from IoT devices (Internet of Things) or monitoring systems; and,

· Medical insurance claims.

Secure information sharing methods, which allow both healthcare providers and their covered entities to verify the correctness of data, are crucial for ensuring proper medical services. This is where blockchain comes in useful, as one of its main advantages is data integrity. When information is recorded and encrypted, it becomes impossible to change or remove.

One of the blockchain approaches that allows for the secure recording and sharing of information is anchoring data to the public blockchain. This method involves generating a proof of data integrity. Using this proof, any user can verify the data timestamp without the need to rely on third-parties. This method allows users to:

· Verify PHI integrity;

· Perform unchangeable medical audits;

· Prove the integrity of clinical research results;

· Reduce audit expenses and ensure regulatory compliance; and,

· Ensure data safety.

HIPAA requires the usage of safe methods of communication between those who deal with PHI stored in electronic form. That is why data encryption plays a crucial role in ensuring data privacy and safety. Our team has a deep expertise in developing digital solutions for the healthcare industry. One of our projects is a HIPAA compliant online communication platform called MDChat that allows patients to securely communicate with medical employees and be sure they are protected from any hacker attacks.

More Secure Standards

Blockchain provides a more secure way to protect data than ordinary encryption. The new technology allows for the implementation of new standards in managing insurance claims, PHI, and medical records. It excludes intermediation in data sharing, when using blockchain. Such consortiums as Hyperledger help increase awareness of the advantages of cryptography and further explain how to use blockchain in healthcare.

According to the survey mentioned above, the main reason why medical organizations hesitate to use blockchain is the lack of knowledge around this technology. A quarter of respondents are still at the stage of education and exploration, which is why responsible state organizations should make the corresponding information more widespread among caregivers. Healthcare providers suppose that this technology must pass several milestones before any adoption is possible, including:

· Technical proof of concept (PoC) (65.4%);

· Security proof (38.5%);

· Privacy proof (34.6%); and,

· Regulatory approval (23.1%).

We can spend a lot of time wondering why caregivers hesitate to implement blockchain in their organization, though the answer is far simpler than it may seem: they simply do not know enough about this technology and its advantages.

Data Transparency

Besides disintermediation, data integrity and provenance, healthcare providers see transparency (55.2%) as one of the top advantages of using blockchain in their industry. To better understand this aspect, let’s consider how it works in the financial sector.

This technology provides a decentralized register of ownership by recording every transaction made through the system. It stores all details starting from the formation of a data block, and ending with any digits related to a specific transaction. Every device that is a part of the system stores a copy of this block. Before making a transaction, the system confirms whether a blockchain version coincides with another in the network. Therefore, each blockchain user can identify the owner of a particular data block at any time. Furthermore, the blockchain is not only a secure way to send money, but a fully protected data sharing method that widens its potential use in healthcare.

Blockchain in Healthcare: Usage

Caregivers feel quite optimistic about fast blockchain implementation, with 37.9% predicting that it will take only five years to adopt it across medical organizations. For now, these organizations and professionals need examples of blockchain, and how it can be helpful in their field. Here, we will cover examples of blockchain use in the healthcare industry, describing existing issues in the sector and considering possible solutions through the use of this technology.

Blockchain in healthcare examples include the following usage issues:

· Drug traceability;

· Data security in clinical trials; and,

· Patient Data Management.

Let's consider each of them.

Problem: Drug Traceability

One of the most serious problems in pharmacology is drug counterfeit. According to the Health Research Funding Organization (HRFO), approximately 10%-30% of drugs in developing countries are fake. US businesses lose up to $200 billion annually because of drug counterfeiting; however, the main reason is not in counterfeiting itself, but, rather, that these drugs provide different effects than their traditional medicinal counterparts. They may not help patients at all, or may even be harmful and dangerous to a person's health.

Blockchain-Based Solution

As all transactions in blockchain are times tamped and immutable, it is easy to detect fraudulent drug dealers. There are two blockchain types: private and public. Trustworthy healthcare blockchain companies have to register their products in the private system to ensure authenticity and the high quality of their medicines. Private blockchains are moderated by central entities, and the fact that a specific producer or distributor has access to the so-called drug blockchain is proof of drug authenticity. This is where blockchain transparency comes in useful. Once a drug is produced and moves from the manufacturer to retailer, the operational data is recorded on the blockchain. It makes it extremely easy to verify the whole path of the drug, and determine all chain links at any time.

Problem: Data Security in Clinical Trials

Clinical trials are used to determine the effectiveness of particular medicines which cure specific diseases. These tests can either prove or disprove an offered hypothesis. During clinical trials, researchers obtain and record a great deal of information concerning statistics, test results, quality reports, etc. Each scientist is responsible for specific research, making it difficult to control everyone. Those data can then be easily modified or hidden in order to change the whole outcome of the research performed. Criminals are interested in recording the results that are beneficial for them, even if the data does not coincide with the reality.

Blockchain-Based Solution

This technology allows users to prove the authenticity of any document registered in the system. It provides proof-of-existence by adding data in the form of the transaction and validating the information by all system nodes. As mentioned above, blockchain records immutable data. This characteristic will allow for the storage of results from clinical trials in a secure way, making it impossible to modify data. Two doctors from Cambridge University conducted a 2016 study to see how blockchain can provide proof-of existence for clinical trials. They found that comparing a unique data code, which is set by the system, with the original makes it possible to verify whether the data of clinical trials has been modified, thanks to the inner SHA256 calculator which generates a unique hash every time a modification is made to the data.

Problem: Patient Data Management

Patient data privacy is strictly regulated by the Health Insurance Portability and Accountability Act (HIPAA), and requires PHI to be totally secure. There is, however, another problem related to PHI: sometimes, patients need to share their medical records with third parties (e.g. with pharmacies when they need to buy specific medicines). So, how can blockchain help protect data while providing partial access at the same time?

Blockchain-Based Solution

The Blockchain creates a hash for each PHI block, together with a patient ID. Using an API, covered entities can receive the necessary information without revealing a patient's identity. In the same way, a patient can decide whom to provide with access and whether this access will be either full or partial. Furthermore, a patient can set specific third parties that would have to give their permission for sharing the PHI, if the patient is not sure in what he or she is doing.

Blockchain has a tremendous potential of use in different industries, including healthcare. This technology has already become widespread in the financial sector, but medical organizations still hesitate to implement it into their IT systems. This does not mean, however, that there are no healthcare companies currently using blockchain. Below, you will find a short list of startups that have made this technology the base of their operational structure.

Blockchain is an effective technology that can help prevent data breaches in the healthcare industry. It is a secure and reliable method of recording, storing, and sharing sensitive data. Caregivers will definitely benefit from implementing this technology, while remaining HIPAA compliant with this method of trustworthy digital protection.

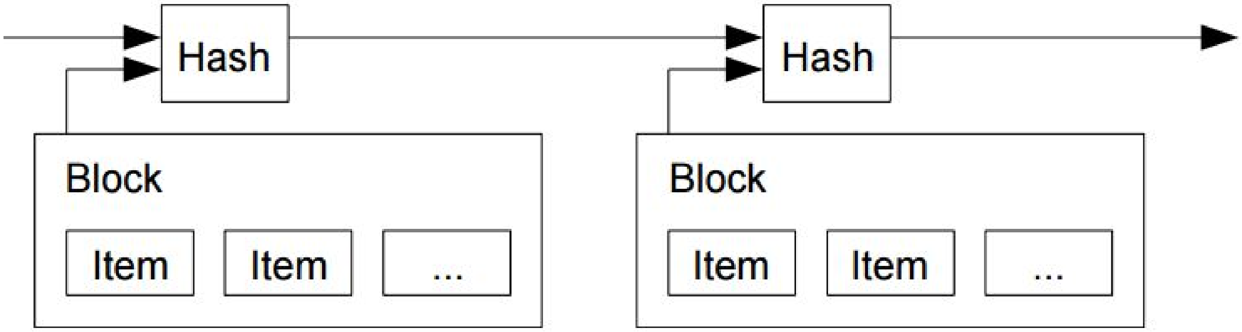
DOCUMENT MANAGEMENT

Document management is the blockchain based application.In this application the user can store the document by uploading it and all can download the document.This application is secure and can manage the different version of a document .Any updation in any document will consider as a new version of this document.As it is a blockchain based application so no one can delete the document.It gives the transparency to all to see and download all the versions of document.

Asymmetric cryptography and distributed consensus algorithms have been implemented for user security and ledger consistency (Zheng et al, 2017). The blockchain technology generally has key characteristics of decentralization, persistency, anonymity and auditability. With these traits, blockchain can greatly save the cost and improve the efficiency.blockchain is immutable. Transaction cannot be tampered once it is packed into the blockchain.

This blockchain application is distributed and can avoid the single point of failure situation.As for smart contracts, the contract could be executed automatically once the contract has been deployed on the blockchain.

The blockchain is a sequence of blocks, which holds a complete list of transaction records like conventional public ledge. Each block points to the immediately previous block via a reference that is essentially a hash value of the previous block called parent block. It is worth noting that uncle blocks (children of the block‘s ancestors) hashes would also be stored in ethereum blockchain (Buterin, 2014). The first block of a blockchain is called genesis block which has no parent block.



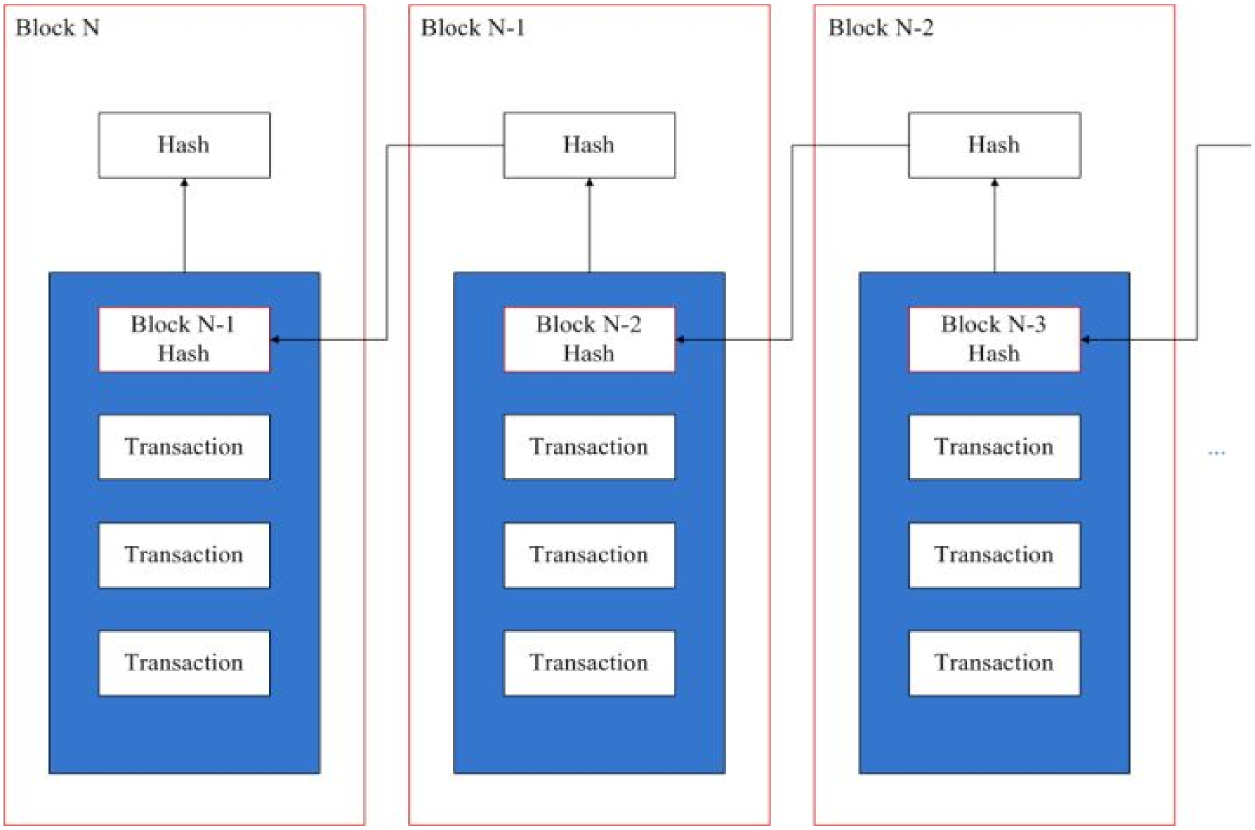


Figure 1 an example of blockchain which consists of a continuous sequence of blocks

Blockchain Block

A block is an aggregated set of data. Data are collected and processed to fit in a block through a process called mining. Each block could be identified using a cryptographic hash (also known as a digital fingerprint). The block formed will contain a hash of the previous block, so that blocks can form a chain from the first block ever (known as the Genesis Block) to the formed block. In this way, all the data could be connected via a linked list structure (Eyal & Sirer, 2018). In particular, the block header includes:
Block version: indicates which set of block validation rules to follow.
Parent block hash: a 256-bit hash value that points to the previous block.
Merkle tree root hash: the hash value of all the transactions in the block.
Timestamp: current timestamp.
nBits: current hashing target in a compact format.
Nonce: a 4-byte field, which usually starts with 0 and increases for every hash calculation.
The block body is composed of a transaction counter and transactions. The maximum number of transactions that a block can contain depends on the block size and the size of each transaction.

Blockchain uses an asymmetric cryptography mechanism to validate the authentication of transactions (Aitzhan & Svetinovic, 2018). A digital signature based on asymmetric cryptography is used in an untrustworthy environment.

Major Features of Blockchain applications :

Decentralization :In most conventional centralised transaction systems, transaction requires to be authorised through the central trusted parties (such as the central bank) inevitably causing the cost and the performance hold-ups at the central servers. Inversely, a transaction in the blockchain network can be conducted between any two peers (P2P) without the authentication by the central agency (Tosh, 2017). In this manner, blockchain can significantly mitigate the performance bottlenecks at the central server. Through this, people trading on a blockchain based applications have a direct control over their accounts by the means of a key that is linked to their accounts which gives the owners a power to transfer their assets to anyone they want. The Blockchain technology has proved to be a really effective tool for decentralizing the web. And it does possess the power to bring massive changes in the industries.

Immutability. Since each of the transactions spreading across the network needs to be confirmed and recorded in blocks distributed in the whole network, it is nearly impossible to tamper. Additionally, each broadcasted block would be validated by other nodes and transactions would be checked. So any falsification could be detected easily.

Anonymity. Each user can interact with the blockchain network with a generated address. Further, a user could generate many addresses to avoid identity exposure. There is no longer any central party keeping users‘ private information. This mechanism preserves a certain amount of privacy on the transactions included in the blockchain. Note that blockchain cannot guarantee the perfect privacy preservation due to the intrinsic constraint that are yet to be discussed.

Auditability. Since each of the transactions on the blockchain is validated and recorded with a timestamp, users can easily verify and trace the previous records through accessing any node in the distributed network (Zheng et al, 2016). In Bitcoin blockchain, each transaction could be traced to previous transactions iteratively (Aitzhan & Svetinovic, 2018). This improves the traceability and the transparency of the data stored in the blockchain.

There mainly three types of Blockchains that have emerged after Bitcoin introduced Blockchain to the world: public blockchain, private blockchain and consortium blockchain (Guo & Liang 2016).

Public blockchain- all records are visible to the public and everyone could take part in the consensus process. Here no one is in charge and anyone can participate in reading/writing/auditing the blockchain (Sato & Matsuo, 2017). These types of blockchain are open and transparent hence anyone can review anything at a given point of time on a public blockchain. For example Bitcoin, Litecoin and ethereum (Gervais et al, 2016). With these blockchain networks.

This application is based on private blockchain :

Private blockchain- This is a private property owned by an individual or an organization. Only those nodes that come from one specific organization would be allowed to join the consensus process Unlike public blockchain this chain has an in charge who looks after of important things such selectively giving access to read or vice versa. Here the consensus is achieved on the whims of the central in-charge who can give mining rights to anyone or not give at all. This makes it centralized again where various rights are exercised and vested in a central trusted party but yet it is cryptographically secured from the company's point of view and more cost-effective for them. For Example: Bankchain. With these types of blockchain networks;

Anybody cannot run a full node and start mining.

Anybody cannot make transactions on the chain.

Anybody cannot review/audit the blockchain in a Blockchain explorer.

Consortium blockchain- This blockchain tries to remove

Blockchain applications: Currently most Blockchains are used in the financial domain, more and more applications for different fields are appearing. Traditional industries could take blockchain into consideration and apply blockchain into their fields to enhance their systems. For example, user reputations could be stored on blockchain. At the same time, the up-and-coming industry could make use of blockchain to improve performance. A smart contract based on blockchain transaction protocol that executes the terms of a contract are also possible. In blockchain, smart contract is a code fragment that could be executed by miners automatically (Luu et al, 2016). Smart contract has transformative potential in various fields li

Technology used

Blockchain :

Blockchain technology is not a company, nor is it an app, but rather an entirely new way of documenting data on the internet. The technology can be used to develop blockchain applications, such as social networks, messengers, games, exchanges, storage platforms, voting systems, prediction markets, online shops and much more.

The information recorded on a blockchain can take on any form, whether it be denoting a transfer of money, ownership, a transaction, someone's identity, an agreement between two parties, or even how much electricity a lightbulb has used. However, to do so requires a confirmation from several of devices, such as computers, on the network. Once an agreement, otherwise known as a consensus, is reached between these devices to store something on a blockchain it is unquestionably there, it cannot be disputed, removed or altered, without the knowledge and permission of those who made that record, as well as the wider community.

Rather than keeping information in one central point, as is done by traditional recording methods, multiple copies of the same data are stored in different locations and on different devices on the network, such as computers or printers. This is known as a peer to peer (P2P) network. This means that even if one point of storage is damaged or lost, multiple copies remain safe and secure elsewhere.

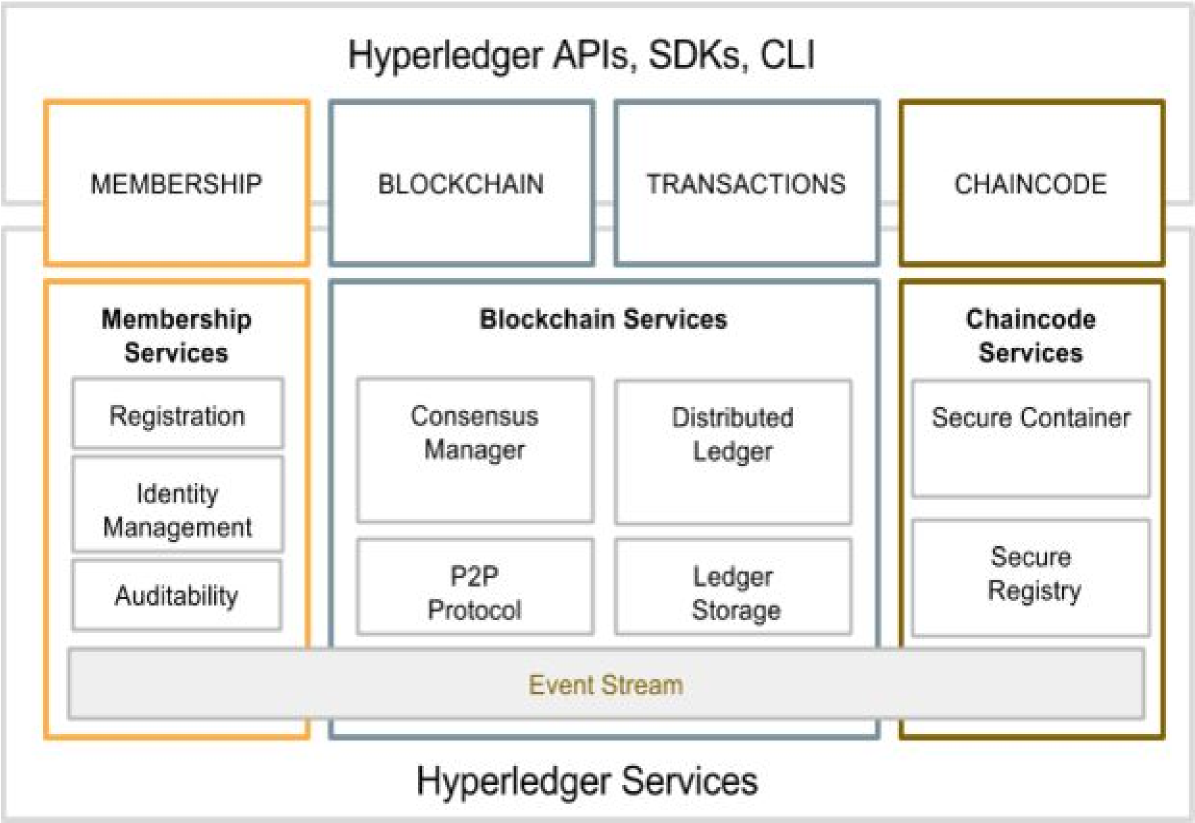
Rather than keeping information in one central point, as is done by traditional recording methods, multiple copies of the same data are stored in different locations and on different devices on the network, such as computers or printers. This is known as a peer to peer (P2P) network. This means that even if one point of storage is damaged or lost, multiple copies remain safe and secure elsewhere.

Hyperledger fabric

“Hyperledger is an open source collaborative effort created to advance cross-industry blockchain technologies. It is a global collaboration, hosted by The Linux Foundation, including leaders in finance, banking, Internet of Things, supply chains, manufacturing, and Technology.”

It is a private and permissioned Blockchain system which means Unlike, in Permissionless(or public network) systems that allow unknown identities to participate in the network, the members enroll through Membership Service Provider (MSP).

It also offers the ability to create channels, allowing a group of participants to create a separate ledger of transactions.



Docker :

Docker is a tool designed to make it easier to create, deploy, and run applications by using containers. Containers allow a developer to package up an application with all of the parts it needs, such as libraries and other dependencies, and ship it all out as one package. By doing so, thanks to the container, the developer can rest assured that the application will run on any other Linux machine regardless of any customized settings that machine might have that could differ from the machine used for writing and testing the code.

Nodejs :

Node.js is a platform built on Chrome's JavaScript runtime for easily building fast and scalable network applications. Node.js uses an event-driven, non-blocking I/O model that makes it lightweight and efficient, perfect for data-intensive real-time applications that run across distributed devices.
Node.js is an open source, cross-platform runtime environment for developing server-side and networking applications. Node.js applications are written in JavaScript, and can be run within the Node.js runtime on OS X, Microsoft Windows, and Linux.

Node.js also provides a rich library of various JavaScript modules which simplifies the development of web applications using Node.js to a great extent.

Go :

Go is an open source programming language that makes it easy to build simple, reliable, and efficient software.
Go is a procedural programming language. It was developed in 2007 by Robert Griesemer, Rob Pike, and Ken Thompson at Google but launched in 2009 as an open source programming language. Programs are assembled by using packages, for efficient management of dependencies. This language also supports environment adopting patterns alike to dynamic languages. For eg., type inference (y := 0 is a valid declaration of a variable y of type float).

Angular :

Angular is a platform that makes it easy to build applications with the web. Angular combines declarative templates, dependency injection, end to end tooling, and integrated best practices to solve development challenges. Angular empowers developers to build applications that live on the web, mobile, or the desktop.

Bootstrap :

Bootstrap is a giant collection of handy, reusable bits of code written in HTML, CSS, and JavaScript. It’s also a front-end development framework that enables developers & designers to quickly build fully responsive websites.

Databases :

- MongoDB: MongoDB is a document-oriented NoSQL database used for high volume data storage. MongoDB is a database which came into light around the mid-2000s. It falls under the category of a NoSQL database.

- CouchDB : CouchDB is an open source database developed by Apache software foundation. The focus is on the ease of use, embracing the web. It is a NoSQL document store database.
  It uses JSON, to store data (documents), java script as its query language to transform the documents, http protocol for api to access the documents, query the indices with the web browser. It is a multi master application released in 2005 and it became an apache project in 2008.

- levelDB : LevelDB is an open-source on-disk key-value store written by Google fellows Jeffrey Dean and Sanjay Ghemawat.Inspired by Bigtable, LevelDB is hosted on GitHub under the New BSD License and has been ported to a variety of Unix-based systems, macOS, Windows, and Android.
  LevelDB is not an SQL database. Like other NoSQL and dbm stores, it does not have a relational data model and it does not support SQL queries. Also, it has no support for indexes. Applications use LevelDB as a library, as it does not provide a server or command-line interface.

5. INFORMATION ABOUT MODULE

1. Hyperledger Fabric

Hyperledger Fabric is a blockchain framework implementation and one of the Hyperledger projects hosted by The Linux Foundation. Intended as a foundation for developing applications or solutions with a modular architecture, Hyperledger Fabric allows components, such as consensus and membership services, to be plug-and-play. Hyperledger Fabric leverages container technology to host smart contracts called "chaincode" that comprise the application logic of the system. Hyperledger Fabric was initially contributed by Digital Asset and IBM, as a result of the first hackathon.

Hyperledger Fabric has become the de-facto standard for enterprise blockchain platforms. It offers a unique approach to consensus that enables performance at scale while preserving privacy to deliver an interoperable network-of-networks. Through open source and open governance, it features innovative new capabilities hardened for use by businesses, ushering in a new era of trust, transparency and accountability.

To meet modern business demands, IBM joined with other companies to collaboratively develop an open source, production-ready, business blockchain framework, called Hyperledger Fabric™, one of the 8 Hyperledger® projects hosted by The Linux Foundation®. Hyperledger Fabric supports distributed ledger solutions on permissioned networks for a wide range of industries.

Permissioned network

Establish decentralized trust in a network of known participants rather than a public network with no identity

Confidential transactions

Expose only the data you want to share to the parties you want to share it with

Pluggable architecture

Tailor the blockchain to industry needs with a pluggable architecture rather than a one size fits all approach

Easy to Get Started

Program smart contracts in the languages your team works in today instead of learning custom languages and architectures

1.Permissioned membership

Hyperledger Fabric is a framework for *permissioned networks*, where all participants have known identities. When considering a permissioned network, you should think about whether your blockchain use case needs to comply with data protection regulations. Many use cases — in the financial sector and healthcare industry, in particular — are subject to data protection laws that require knowing who the members of the network are and who is accessing specific data.

For example, consider a private equity company. By definition, a private equity is not publicly traded on the stock exchange, and its investors are typically venture capital firms, private equity firms, or angel investors. The participants in this network need to be known and have credibility in capital to invest to be able to participate in the blockchain.

2. Performance, scalability, and levels of trust

Hyperledger Fabric is built on a modular architecture that separates transaction processing into three phases: distributed logic processing and agreement ("chaincode"), transaction ordering, and transaction validation and commitment. This separation confers several advantages: Fewer levels of trust and verification are required across node types, and network scalability and performance are optimized.

Starting at the left of the figure:

1) The transaction proposal is submitted by an application to an endorsing peer.

2) The Endorsement policies outline how many and/or what combination of endorsers are required to sign a proposal. The endorser executes the chaincode to simulate the proposal in the network peer, creating a read/write set.

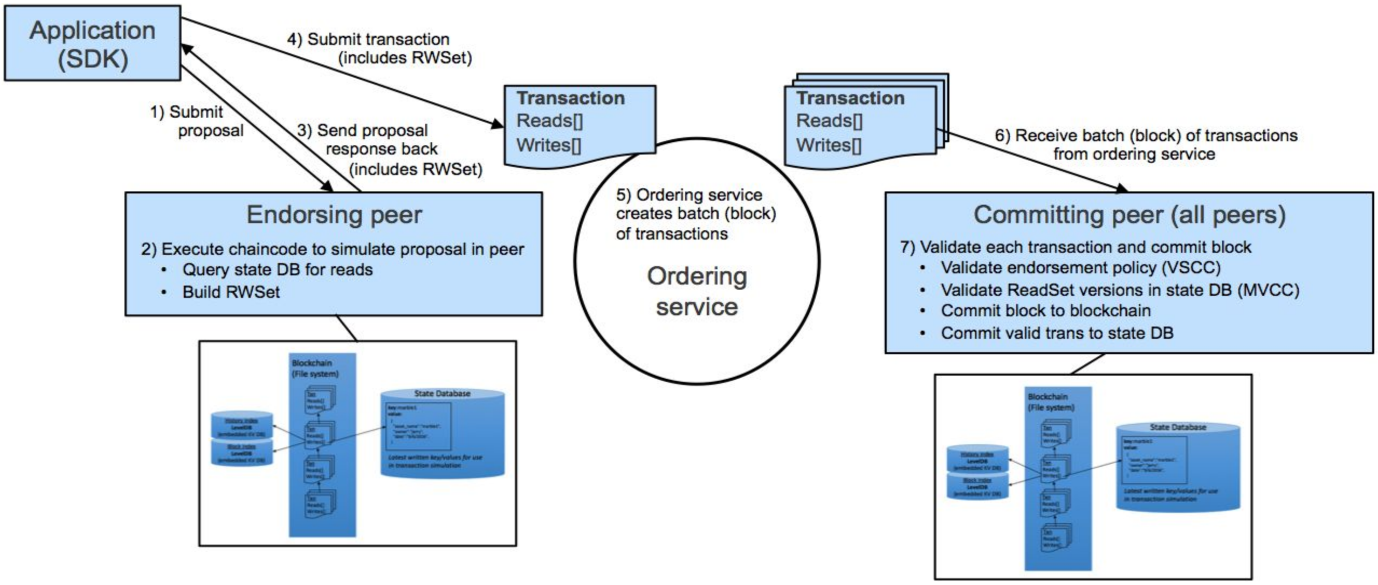
3)Then the endorsing peers send back the signed proposal responses (endorsements) to the application.

4) The application submits the transactions and signatures to the ordering service, which

5) creates a batch, or block, of transactions and delivers them to committing peers.

6) When a committing peer receives a batch of transactions, for each transaction it

7) validates that the endorsement policy was met and checks in the read/write sets to detect conflicting transactions. If both checks pass, the block is committed to the ledger, and the state updates for each transaction are reflected in the state database.


Application (SDK)
1) Submit proposal
3) Send proposal response back (includes RWSet)
4) Submit transaction (includes RWSet)
Transaction
Reads[]
Writes[]
Endorsing peer
2) Execute chaincode to simulate proposal in peer
• Query state DB for reads
• Build RWSet
Blockchain (File system)
State Database
Latest written key/values for use in transaction simulation
5) Ordering service creates batch (block) of transactions
Ordering service
Transaction
Reads[]
Writes[]
6) Receive batch (block) of transactions from ordering service
Committing peer (all peers)
7) Validate each transaction and commit block
• Validate endorsement policy (VSCC)
• Validate ReadSet versions in state DB (MVCC)
• Commit block to blockchain
• Commit valid trans to state DB
Blockchain (File system)
State Database
Latest written key/values for use in transaction simulation

Because only the signatures and read/write set are being sent around the network with the new v1.0 architecture, the scalability and performance are optimized. In addition, because only the endorsers and the committers truly see the transaction, fewer levels of trust are required in different parts of the Blockchain system, offering more security.

For example, in the capital market, with equity-backed securities or debt being bought and sold, transaction volume has increased, due to a growing number of participants. Increased transactions require improved scalability and performance, which v1.0 of Hyperledger Fabric provides, due in part to splitting out the chaincode execution.

Splitting out the chaincode execution also enables dynamic growth in the network. In v1.0 of Hyperledger Fabric, peers can be added dynamically and programmatically, rather than statically as in v0.6. For example, suppose a company that manages foreign exchange rates has a new bank to add to the network. With Hyperledger Fabric v1.0, they can do this programmatically.

3. Data on a need-to-know basis

Businesses, due to competitiveness, protection laws, and regulation on confidentiality of personal data dictate the need for privacy of certain data elements, which can be achieved through data partitioning on the blockchain. Channels, supported in Hyperledger Fabric, allow for data to go to only the parties that need to know.

For example, many financial entities express concern over competitors seeing even the number of transactions being processed. Some financial institutions do not consider cryptography "enough" to protect their data. Given that some financial instruments can take 10 years or more to come to value, the risk of cryptography breaks over time could allow private information to become public. Channels help provide a data-partitioning capability where only those that need to know the data will see the number of transactions and the data itself.

4. Rich queries over an immutable distributed ledger

The ledger is the sequenced record of state transitions for the blockchain application. Each transaction results in a set of asset key-value pairs that are committed to the ledger as creates, updates, or deletes. The immutable source of truth for v1.0 is appended into the file system of the peer, which also has LevelDB embedded.

LevelDB has, by default, a key value database and supports keyed queries, composite key queries, and key range queries. If you also need complex, rich queries, CouchDB supports the basic capabilities of LevelDB, and adds the full data-rich queries. With optional support of a document database such as CouchDB, the content is JSON and fully queryable, where the data model is compatible with existing key/value programming model. As a result, the application changes are not required when modeling chaincode data as JSON when utilizing CouchDB.

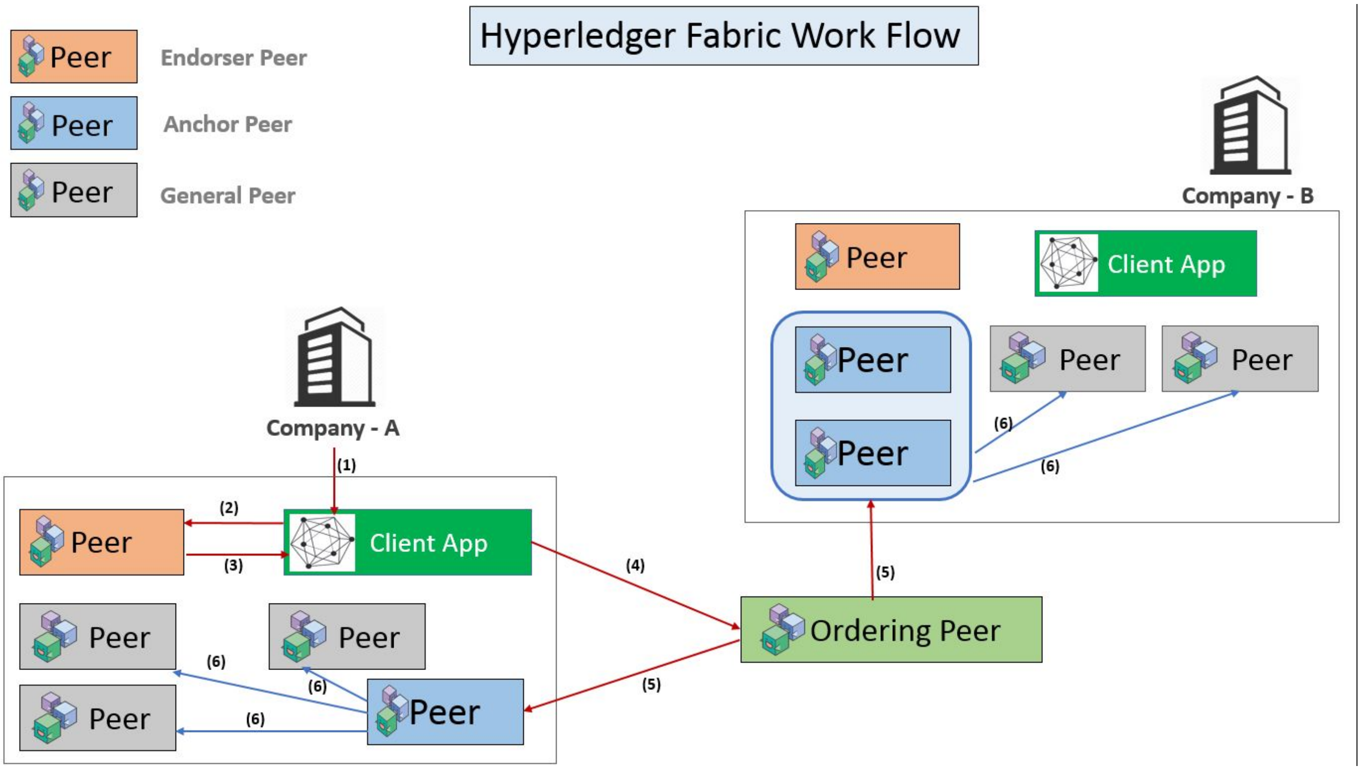
This JSON format helps minimize the work required to produce simple reports and perform audit functions. For example, in supply-chain scenarios, you can use JSON document style to help outline specific data for goods and transportation entities. You can easily produce a report on an asset for the different locations and transportation entities that were used in delivery to the asset's final destination.

5. Modular architecture supporting plug-in components

The modularity of Hyperledger Fabric architecture enables network designers to plug in their preferred implementations for components, which is an advantage. One of the most requested areas for modularity is "bring your own identity." Some multi-company networks already have identity management and want to reuse instead of rebuild. Other components of the architecture that can be easily plugged in include consensus or encryption, where some countries have their own encryption standards.

6. Protection of digital keys and sensitive data

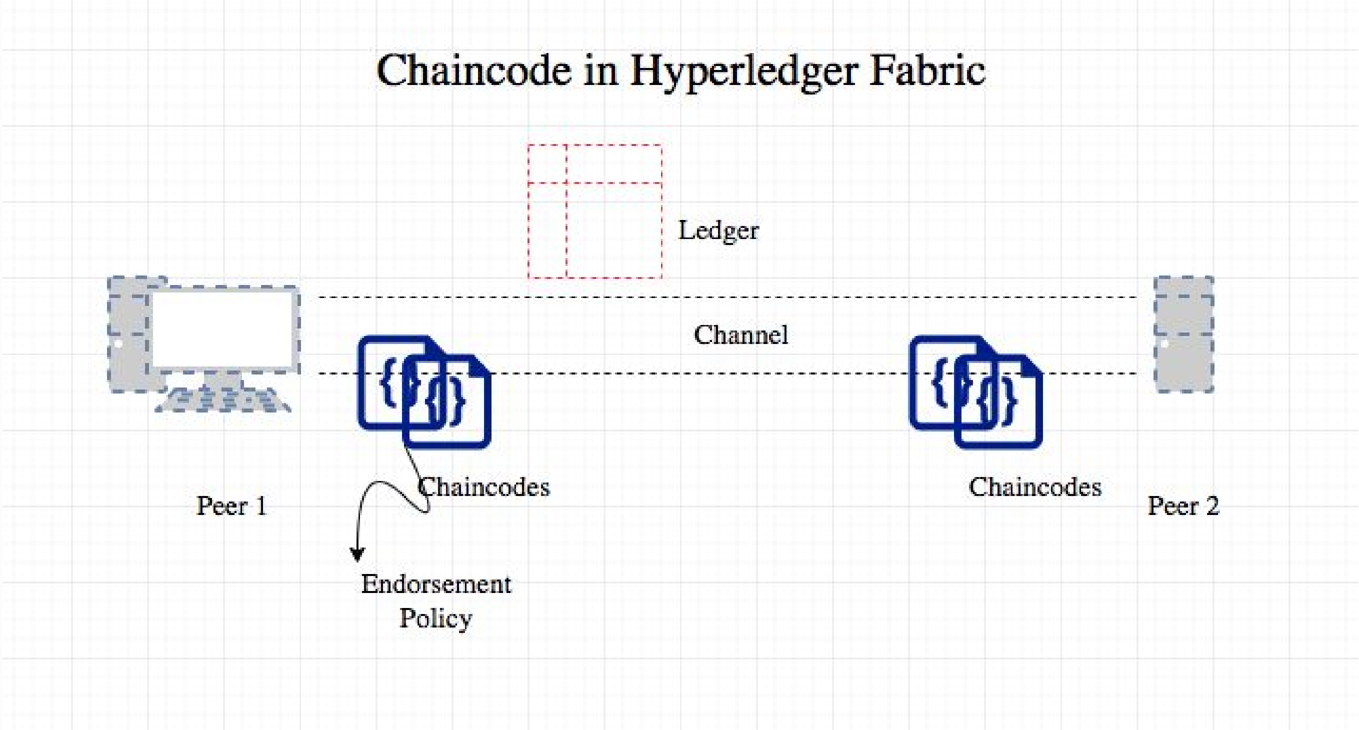
HSM (Hardware Security Module) support is vital for safeguarding and managing digital keys for strong authentication. Hyperledger Fabric provides modified and unmodified PKCS11 for key generation, which supports cases like identity management that need more protection. For scenarios dealing with identity management, HSM increases the protection of keys and sensitive data.



2. Chaincode

Chaincode is a program, written in Go, node.js, or Java that implements a prescribed interface. Chaincode runs in a secured Docker container isolated from the endorsing peer process. Chaincode initializes and manages ledger state through transactions submitted by applications.

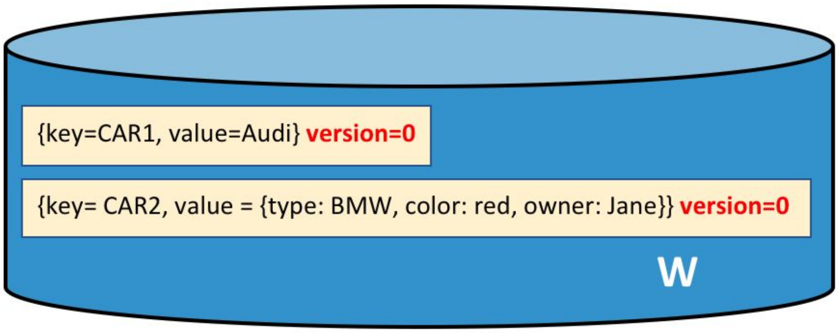
A chaincode typically handles business logic agreed to by members of the network, so it may be considered as a “smart contract”. State created by a chaincode is scoped exclusively to that chaincode and can’t be accessed directly by another chaincode. However, within the same network, given the appropriate permission a chaincode may invoke another chaincode to access its state.



3. CouchDB

Hyperledger Fabric (HLF) uses a key value database to store its state. This object store holds binary data which can be queried using its key. By default fabric uses a LevelDB store which is included in the peer process.

CouchDB is another type of key value store that can be easily plugged into Fabric. Both LevelDB & CouchDB can store binary data and can be interacted with using the chaincode. But CouchDB also adds rich querying to the mix, this of course if you are storing json documents. The "*first network*" sample project also contains a configuration with CouchDB included. The configuration for this can be found on here on the HLF wiki.


{key=CAR1, value=Audi} version=0
{key= CAR2, value = {type: BMW, color: red, owner: Jane}} version=0
W

W	Ledger world state
{key=**K**, value = **V** } **version=0**	A ledger state with **key=K**. It contains a set of facts expressed as a simple value, **V**. The state is at version 0.
{key=**K**, value = {**KV**} } **version=0**	A ledger state with **key=K**. It contains a set of facts expressed as a set of key-value pairs {**KV**}**.** The state is at version 0.

CouchDB feels like a key value store, with the querying ability of MongoDB. While the HLF team has improved their documentation since 1.0-alpha.

Why CouchDB?

Fabric supports two types of peer databases. LevelDB is the default state database embedded in the peer node and stores chaincode data as simple key-value pairs and supports key, key range, and composite key queries only. CouchDB is an optional alternate state database that supports rich queries when chaincode data values are modeled as JSON. Rich queries are more flexible and efficient against large indexed data stores, when you want to query the actual data value content rather than the keys. CouchDB is a JSON document datastore rather than a pure key-value store therefore enabling indexing of the contents of the documents in the database.

In order to leverage the benefits of CouchDB, namely content-based JSON queries,your data must be modeled in JSON format. You must decide whether to use LevelDB or CouchDB before setting up your network. Switching a peer from using LevelDB to CouchDB is not supported due to data compatibility issues. All peers on the network must use the same database type. If you have a mix of JSON and binary data values, you can still use CouchDB, however the binary values can only be queried based on key, key range, and composite key queries.

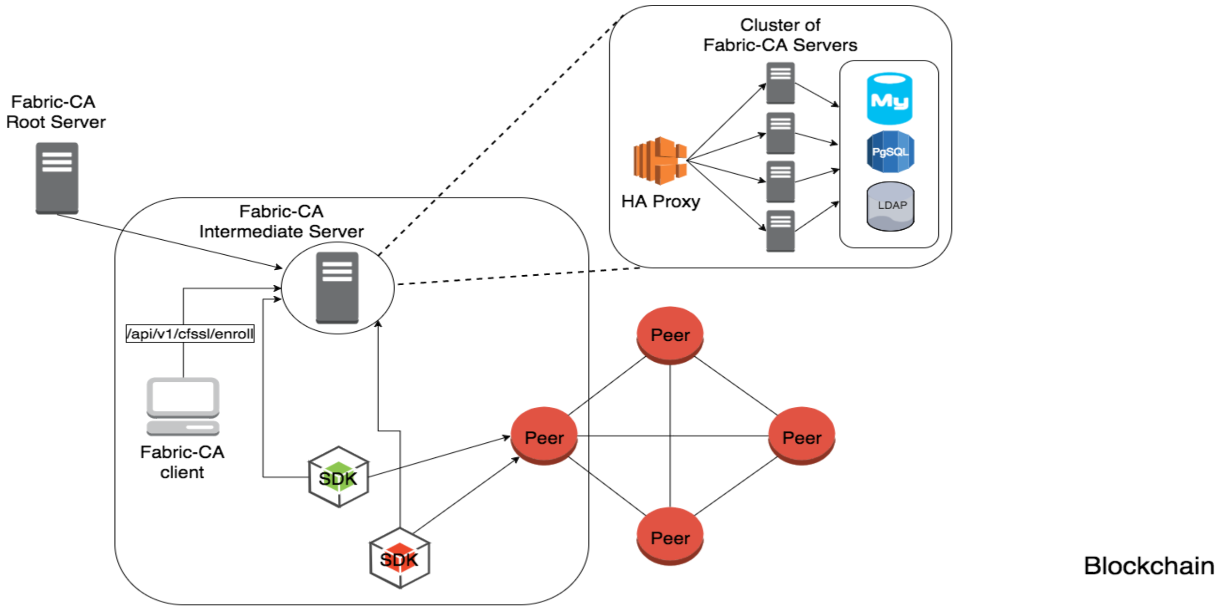
4. Docker

Docker is an open source software platform to create, deploy and manage virtualized application containers on a common operating system (OS), with an ecosystem of allied tools. Docker Inc., the company that originally developed Docker, supports a commercial edition and is the principal sponsor of the open source tool.

Docker is a tool that packages, provisions and runs containers independent of the OS. Container technology is available through the operating system: A container packages the application service or function with all of the libraries, configuration files, dependencies and other necessary parts to operate. Each container shares the services of one underlying operating system.

Docker was created to work on the Linux platform, but has extended to offer greater support for non-Linux operating systems, including Microsoft Windows and Apple OS X. Versions of Docker for Amazon Web Services (AWS) and Microsoft Azure are available.

Docker has emerged as a de facto standard platform that allows users to quickly compose, create, deploy, scale and oversee containers across Docker hosts. Docker allows a high degree of portability so that users can register and share containers over various hosts in private and public environments. Docker benefits include efficient application development, lower resource use and faster deployment compared to VMs.

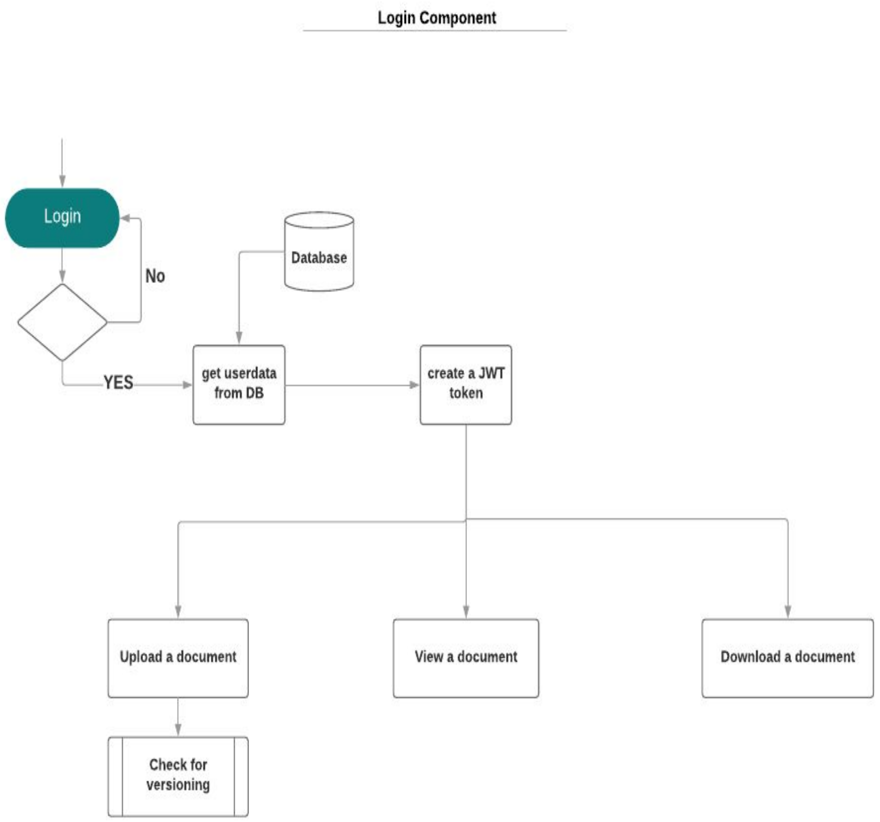


5. Docker Compose

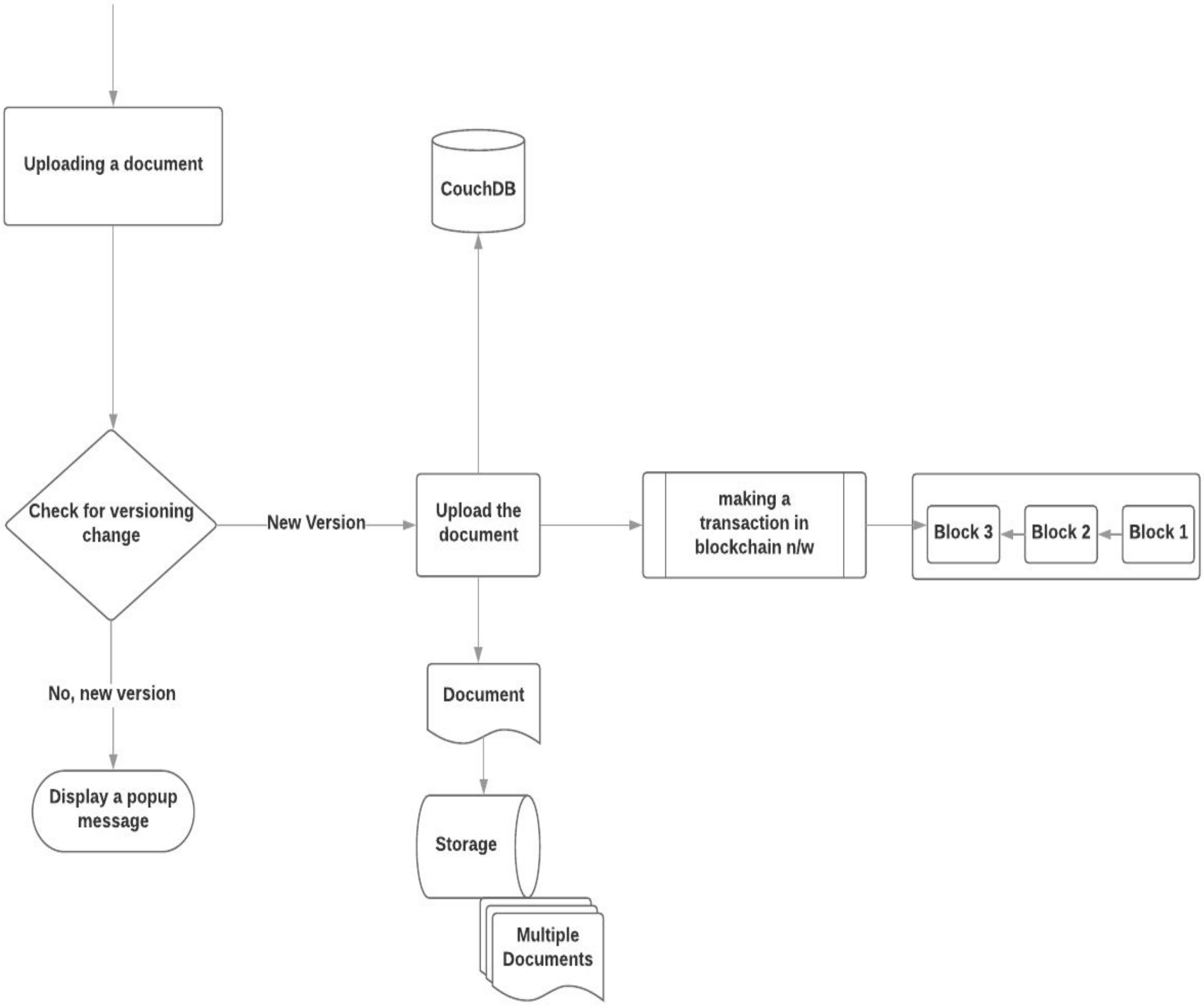
If the Docker application includes more than one container (for example, a webserver and database running in separate containers), building, running, and connecting the containers from separate Dockerfiles is cumbersome and time-consuming. Docker Compose solves this problem by allowing you to use a YAML file to define multi-container apps. You can configure as many containers as you want, how they should be built and connected, and where data should be stored. When the YAML file is complete, we can run a single command to build, run, and configure all of the containers.

6. Design (HLD and LLD UML, DFD etc)

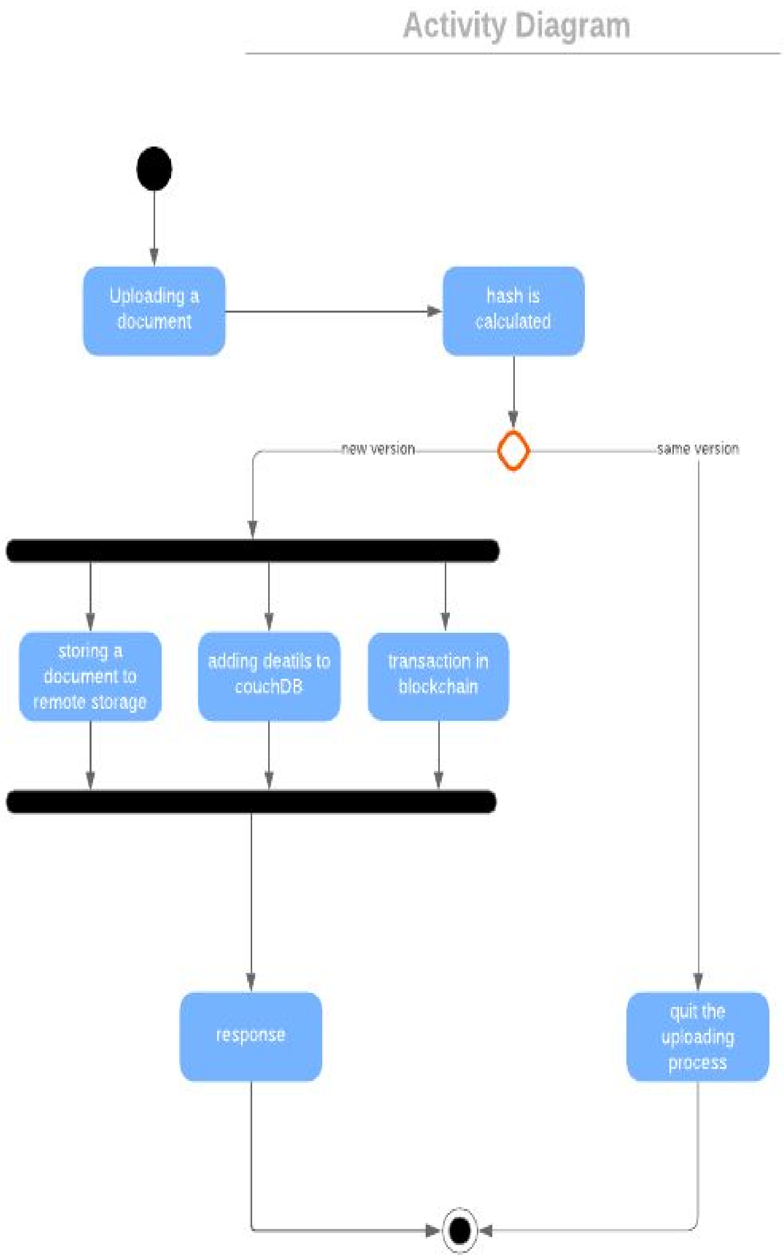
Flowchart:



Upload Component



UML - Activity Diagram


Activity Diagram
Uploading a document
hash is calculated
new version
same version
storing a document to remote storage
adding deatils to couchDB
transaction in blockchain
response
quit the uploading process

7. DATABASE AND DATA DICTIONARY

Mongodb

MongoDB is a document-oriented NoSQL database used for high volume data storage. MongoDB is a database which came into light around the mid-2000s. It falls under the category of a NoSQL database.
MongoDB was developed by **Eliot Horowitz** and **Dwight Merriman** in the year **2007**, when they experienced some scalability issues with the relational database while developing enterprise web applications at their company **DoubleClick**. According to Dwight Merriman, one of the developers of MongoDB, this name of the database was derived from the word *humongous* to support the idea of processing large amount of data.

In 2009, MongoDB was made as an open source project, while the company offered commercial support services. Many companies started using MongoDB for its amazing features. The New York Times newspaper used MongoDB to build a web based application to submit the photos. In 2013, the company was officially named to **MongoDB Inc**.

MongoDB is a cross-platform, document oriented database that provides, high performance, high availability, and easy scalability. MongoDB works on concept of collection and document.
Each database contains collections which in turn contains documents. Each document can be different with a varying number of fields. The size and content of each document can be different from each other.
Below are the few of the reasons as to why one should start using MongoDB

1. Document-oriented – Since MongoDB is a NoSQL type database, instead of having data in a relational type format, it stores the data in documents. This makes MongoDB very flexible and adaptable to real business world situation and requirements.
2. Ad hoc queries - MongoDB supports searching by field, range queries, and regular expression searches. Queries can be made to return specific fields within documents.
3. Indexing - Indexes can be created to improve the performance of searches within MongoDB. Any field in a MongoDB document can be indexed.
4. Replication - MongoDB can provide high availability with replica sets. A replica set consists of two or more mongo DB instances. Each replica set member may act in the role of the primary or secondary replica at any time. The primary replica is the main server which interacts with the client and performs all the read/write operations. The

Secondary replicas maintain a copy of the data of the primary using built-in replication. When a primary replica fails, the replica set automatically switches over to the secondary and then it becomes the primary server.

5. Load balancing - MongoDB uses the concept of sharding to scale horizontally by splitting data across multiple MongoDB instances. MongoDB can run over multiple servers, balancing the load and/or duplicating data to keep the system up and running in case of hardware failure.

CouchDB

CouchDB is an open source database developed by Apache software foundation. It is written in Erlang programming language. The focus is on the ease of use, embracing the web. It is a NoSQL document store database.

It uses JSON, to store data (documents), java script as its query language to transform the documents, http protocol for api to access the documents, query the indices with the web browser. It is a multi master application released in 2005 and it became an apache project in 2008.

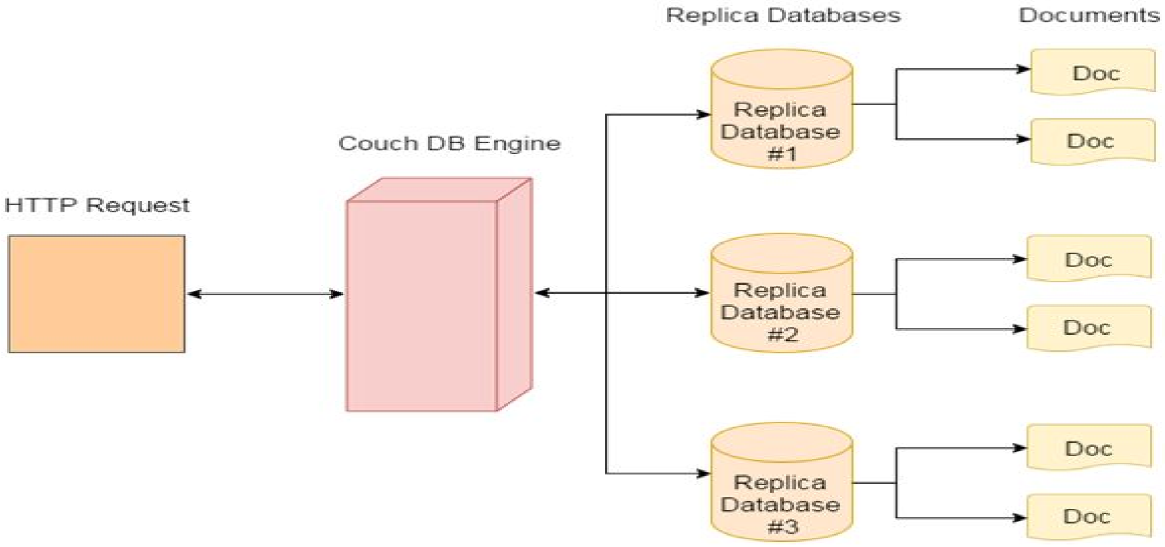
- CouchDB have an HTTP-based REST API, which makes communication with the database very easy.
- CouchDB has the simple structure of HTTP resources and methods (GET, PUT, DELETE) that are easy to understand and use.
- In CouchDB, data is stored in the flexible document-based structure so, there is no need to worry about the structure of the data.
- CouchDB facilitates users with powerful data mapping, which allows querying, combining, and filtering the information.
- CouchDB provides easy-to-use replication, using which you can copy, share, and synchronize the data between databases and machines.

Data Model

- Database is the outermost data structure/container in CouchDB.
- Each database is a collection of independent documents.
- Each document maintains its own data and self-contained schema.
- Document metadata contains revision information, which makes it possible to merge the differences occurred while the databases were disconnected.

- CouchDB implements multi version concurrency control, to avoid the need to lock the database field during writes.

Couchdb Architecture



LevelDB

LevelDB is an open-source on-disk key-value store written by Google fellows Jeffrey Dean and Sanjay Ghemawat. Inspired by Bigtable, LevelDB is hosted on GitHub under the New BSD License and has been ported to a variety of Unix-based systems, macOS, Windows, and Android.

LevelDB stores keys and values in arbitrary byte arrays, and data is sorted by key. It supports batching writes, forward and backward iteration, and compression of the data via Google's Snappy compression library.

LevelDB is not an SQL database. Like other NoSQL and dbm stores, it does not have a relational data model and it does not support SQL queries. Also, it has no support for indexes. Applications use LevelDB as a library, as it does not provide a server or command-line interface.

MariaDB 10.0 comes with a storage engine which allows users to query LevelDB tables from MariaDB.

LevelDB is used as the backend database for Google Chrome's IndexedDB and is one of the supported backends for Riak. Additionally, Bitcoin Core and go-ethereum stores the blockchain metadata using a LevelDB database

Google has provided benchmarks comparing LevelDB's performance to SQLite and Kyoto Cabinet in different scenarios.[11] LevelDB outperforms both SQLite and Kyoto Cabinet in write operations and sequential-order read operations. LevelDB also excels at batch writes, but is slower than SQLite when dealing with large values. The currently published benchmarks were updated after SQLite configuration mistakes were noted in an earlier version of the results. Updated benchmarks show that LevelDB also outperforms Berkeley DB.

Data Dictionary

Basically there are two type of databases used in this project - mongodb for registration and login , couchdb for storing world state of the blockchain application.Both these databases are type nosql that's why the table corrospond to collection.

Mongodb Data dictionary

Number of collection =1

Collection's field or key name = name , email, username ,password.

Key name	Data type	size(maximum)	description
Name	String	4 MB	Name of the person who has registered
Email	String	4 MB	Email of the person who has registered
Username	String	4 MB	Username choosed by the person who has registered
Password	String	4 MB	Password set by the person at the registration time

Couchdb data dictionary

No of Fileds = 5

Name of collection = Upload

Name of fields = name, date, author ,hash , mimeType, location .

Field Name	Data Type	size	Description
Author	String	The total length of the sequence can be up to 2,147,483,648.	Author of the document
Filename	String	The total length of the sequence can be up to 2,147,483,648.	Name of the File to be uploaded
Date	String	The total length of the sequence can be up to 2,147,483,648.	Date and time when the file was uploaded
Type	String	The total length of the sequence can be up to 2,147,483,648.	Type of file i.e, doc,pdf,png etc
Hash	String	The total length of the sequence can be up to 2,147,483,648.	Hash of the file to be stored in chaincode