Blockchain based resilient data storage

Abhishek Bhardwaj
North Carolina State University
abhardw3@ncsu.edu

Madhu Vamsi Kalyan Machavarapu
North Carolina State University
mmachav@ncsu.edu

Sanya Kathuria
North Carolina State University
skathur2@ncsu.edu

Shivam Chamoli
North Carolina State University
schamol@ncsu.edu

ABSTRACT

The Digital world has produced a variety of new innovative products, and close customer relationships globally by the efficient use of mobile, IoT (Internet of Things), social media, analytics and cloud technology to generate models for better decisions. Blockchain was recently introduced and revolutionized the digital world bringing a new perspective to security, resilience, and efficiency of systems. While initially popularized by Bitcoin, Blockchain is much more than a foundation for cryptocurrency. It offers a secure way to exchange any kind of good, service, or transaction. Secure in a way that to create an authorized block, the block needs to be mined with some stringent conditions on how the Hash of the block is created and has to be passed by more than 50% of the peer to peer network, making it extremely resilient to change. Moreover, Blockchain will enable more agile value chains, faster product innovations, closer customer relationships, and quicker integration due to its distributed, decentralized nature.

1. INTRODUCTION

A blockchain is essentially a distributed database of records or a public ledger of all transactions or digital events that have been executed and shared among the participating parties. Each transaction in the public ledger is verified by the consensus of a majority of the participants in the system and, once entered, information can never be erased. The main hypothesis is that the blockchain establishes a system of creating a distributed consensus âĂŃin the digital online world. With public health, wealth, safety, security and environmental protection as a priority, licensing bestows accountability and liability to those developing and operating digital systems and privacy records. We, in this project, would like to address the problem of frauds that take place due to forgery of documents that are present with for example the rotary (like property papers). Another issue that we would like to address is to remove the governing authorities from the use cases like betting, where two parties have to trust a third-party to come through on the deal. Our goal is to make a distributed, decentralized record of potentially fraudulent documents or activities but due to the time-constraints, making a truly generic system which would authenticate the information is unrealistic and hence we would be producing a simple Proof of Concept

X-XXXXX-XX-X/XX/XX.
DOI:

2. MOTIVATION

2.1 Why not Traditional databases?

Traditional databases are usually maintained by a single organization, and that organization has complete control of the database which includes the ability to tamper with the stored data, to censor otherwise valid changes to the data, or to add data fraudulently. Even if we assume that the responsible organization would never enact a fraudulent change to the database which is already too much to ask, there is still the possibility that a hacker could break in and manipulate the database to their own ends.

2.2 Why Blockchain based technology?

The main motivation for this project is to implement a system that makes the database public and allows anyone to store a redundant copy of the database. This is the best way to ensure that data is safely stored.

2.3 How a Blockchain technology solves the problem?

Blockchain technology solves these problems by creating a network of computers (called nodes) which each store a copy of the database, and a set of rules (called the consensus protocol) which define the order in which nodes may take turns adding new changes to the database. In this way, all of the nodes agree to the state of the database at any time, and no one node has the power to falsify the data or to censor changes. The blockchain further requires that an audit trail of all changes to the database is preserved, which allows anyone to audit that the database is correct at any time. This audit trail is composed of the individual changes to the database, which are called transactions.

2.4 Properties of our blockchain

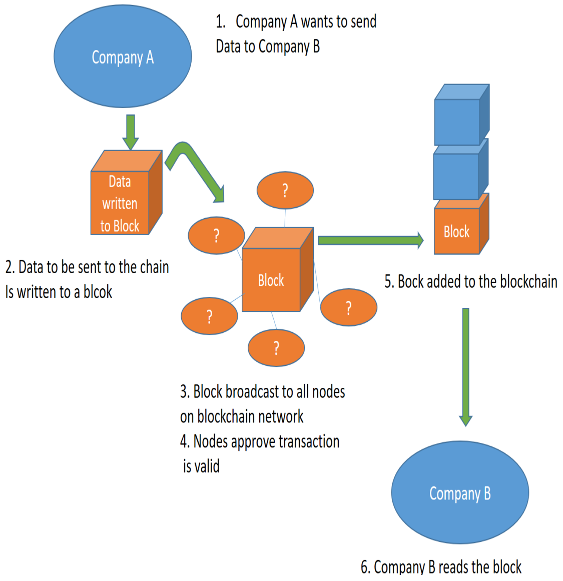
Company A
1. Company A wants to send Data to Company B
Data written to Block
2. Data to be sent to the chain Is written to a blcok
?
?
?
?
?
Block
3. Block broadcast to all nodes on blockchain network
4. Nodes approve transaction is valid
Block
5. Bock added to the blockchain
Company B
6. Company B reads the block

Figure 1: How does Blockchain technology work?

In this short time frame, we could come up with a blockchain constituting of the following properties

2.4.1 *Replication*

Data on the blockchain is copied on every computer that is a part of the P2P network.

2.4.2 *No Central Authority*

There is no central body which governs whether a particular transaction should be recorded or not. This is solved by using consensus amongst all the nodes on the P2P network.

2.4.3 *Irreversibility*

If data/transaction is recorded once on the blockchain, it is very tough to be reversed or updated.

2.4.4 *Accessibility*

Blockchain allows everybody who is a part of the network to read or view any data recorded by anybody.

2.4.5 *Time-stamping*

Timestamping is the process of securely keeping track of the creation and modification time of a document. It allows interested parties to know, without a doubt, that a document in question existed at a particular date and time. In blockchain, every block is time-stamped.

2.5 User Surveys

Please refer to figure 2 and figure 3. The graphs are based on 27 responses.

2.6 Use-case Study

The advantages of Blockchain technology outweigh the regulatory issues and technical challenges. One key emerging use case of blockchain technology involves Digital Blockchain based Property. Digital Blockchain based Property will basically be a computer programs that can automatically execute the terms of a contract. When a pre-configured condition in a smart contract among participating entities is

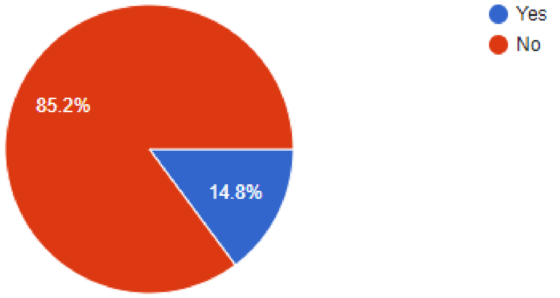

Figure 2: Have you or your friend been duped by forged documents relating to property or shares?

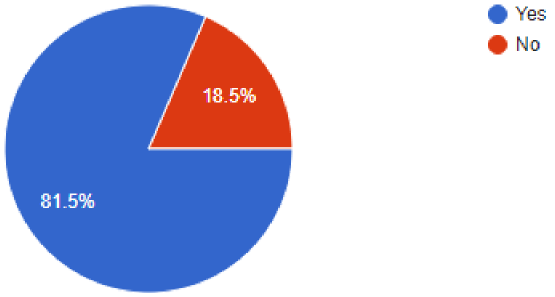

Figure 3: Would you consider a blockchain based online public storage for important documents that is hacker and forgery proof?

met then the parties involved in a contractual agreement can be automatically made payments as per the contract in a transparent manner. Fraud-resistant data keeping of sensitive information.

2.6.1 *Records of property*

Maintaining a national register of property ownership is an expensive and labor-intensive operation. Additionally, in countries where there is a history of government corruption, they may not always be trustworthy. Property documents can be forged by miscreants for stealing properties. With blockchain based data storage this can be made really difficult to execute because of the security features baked inherently in the blockchain.

2.6.2 *Vehicle renting*

The vehicle rental process is often more cumbersome than it needs to be, with insurance documents and identities that need to be verified, and vehicle mileage and damage reports that are still manually verified in many cases. Our project involves implementing a blockchain based data storage such that these cumbersome tasks would be circumvented.

2.6.3 *Information of Stock markets*

The item can be non-physical such as shares of a company.

2.6.4 *Smart Property*

Smart Property is another related concept which is regarding controlling the ownership of a property or asset via blockchain using Smart Contracts. The property can be physical such as car, house, smartphone etc. or it can be non-physical such as shares of a company. It should be noted here that even Bitcoin is not really a currency- Bitcoin is all about controlling the ownership of money.

2.7 Advantages of the system according to our use cases

1. Conflict-proof distributed ledger of probable fraudulent transactions like property transfer, stocks etc.

2. Redundancy of sensitive data because of the distributed and decentralized nature.

3. Fraud - proof, because, to change the value of an already existing block in the blockchain, the person has to recompute all the blocks again and gain access of more than 50 % of the p2p network and recompute the nodes again.

4. Further Blockchain provides a lower cost of trade with a trusted contract monitored without intervention from third parties who may not add direct value.

2.8 Disadvantages of the system according to our use cases

Although the blockchain technology has great potential for the construction of the future Internet systems, it is facing a number of technical challenges.

1. People might not be comfortable with such sensitive data online.

2. The block size is limited to 1 MB now while a block is mined about every ten minutes, which implies that the network is restricted to a rate of 7 transactions per second, which makes it incapable of dealing with high-frequency work. But also, larger blocks means larger storage space and slower propagation in the network. This will lead to centralization gradually as fewer users would like to maintain such a large blockchain. Therefore the tradeoff between block size and security will be a tough challenge.

3. Furthermore, current consensus algorithms like proof of work or proof of stake are facing some serious problems. For example, proof of work wastes too much electrical energy while the phenomenon that the rich get richer could appear in the proof of stake consensus process.

4. Highly unlikely, but if the peer-to-peer network is brought down, all the information is compromised, to handle that we need physical records.

3. EVALUATION PLAN

1. Firstly, we will have to create the front-end which would interact with the user (restrict the inputs) so that our mildly generic blockchain application can handle some cases seamlessly.

2. Building the algorithm for adding blocks to the blockchain.

3. Build a Consensus Algorithm, which is required to accept a block in the transaction list.

4. Setup a peer to peer network and publish the consensus algorithm on each node. Each node would have the transaction list too.

5. Test the security features of the blockchain.

6. Test everything.

7. Repeat 1 - 7 (agile) for improvement.

4. EXECUTION

The execution phase for coding includes (Please refer figure 4)

1. What is an application without user-interaction? Hence we have our novel front-end which would take care of all the user information validation and create url to call a suitable controller.

2. Controller in-turn runs server side validations and forwards the information.

3. The information is run through a consensus algorithm and depending on the validation of the block, the block is either added or the user is redirected to the front-end.

4. If the block is accepted, All the lists in the P2P network are updated and the user is shown the updated list. (List is shown for project purpose, logically, it makes no sense to show the list to the user).

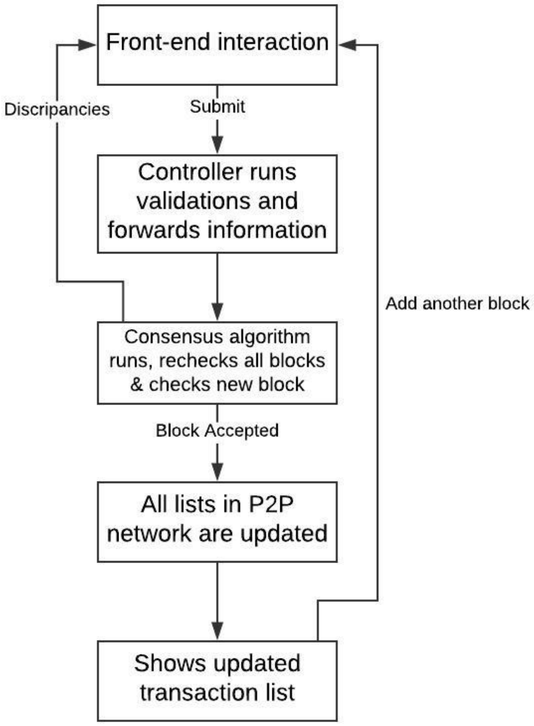
Front-end interaction
Submit
Discripancies
Controller runs validations and forwards information
Add another block
Consensus algorithm runs, rechecks all blocks & checks new block
Block Accepted
All lists in P2P network are updated
Shows updated transaction list

Figure 4: Execution Flow for our Blockchain technology?

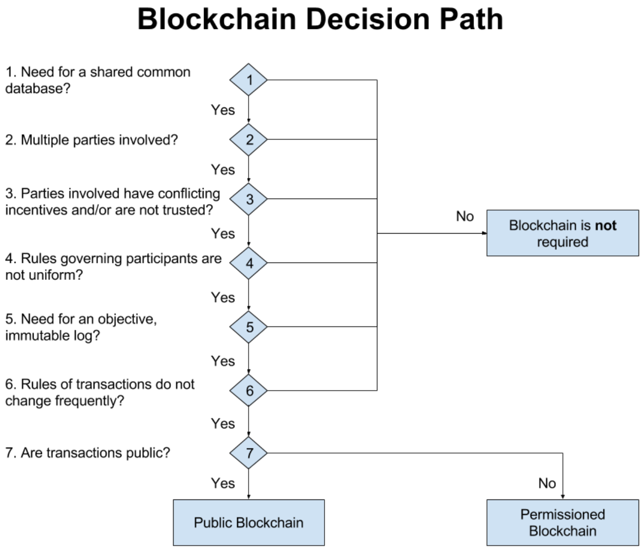
Blockchain Decision Path
1. Need for a shared common database?
Yes
2. Multiple parties involved?
Yes
3. Parties involved have conflicting incentives and/or are not trusted?
Yes
4. Rules governing participants are not uniform?
Yes
5. Need for an objective, immutable log?
Yes
6. Rules of transactions do not change frequently?
Yes
7. Are transactions public?
Yes
No
No
1
2
3
4
5
6
7
Blockchain is not required
Public Blockchain
Permissioned Blockchain

Figure 5: When to use Blockchain technology?

5. TESTING

Testing is a critical task for a blockchain based storage. When a blockchain is created, it should be immutable, that is once a transaction is deployed onto the blockchain it should stay forever.

5.1 Unit testing

Unit testing of Blockchain data integrity, validating encryption and hashcode computation of each block.

5.2 Front-end testing

It would help us validate methods in our blockchain system. This is essentially similar to API testing where one would use method validations, boundary value analysis, decision tables, test driven development and behavior driven development techniques.

5.3 Integration testing

This involves integration testing which emphasizes on the performance and consistency testing between various nodes in the peer to peer network.

6. POSSIBLE FUTURE DIRECTIONS

The blockchain technology has shown a lot of potential in industry and academia. The following are the possible future directions. There can be main areas of future work combined with Blockchain technology: blockchain testing, big data analytics, and blockchain application.

6.1 Blockchain Testing

Recently different kinds of blockchains have appeared but some developers might falsify their blockchain performance to attract investors driven by the huge profit. So, Blockchain testing mechanism needs to be in place to test different blockchains as users want to combine blockchain into business, they have to know which blockchain fits their requirements. .Blockchain testing could be separated into two categories: standardization phase and testing phase. In standardization phase, all criteria have to be made and agreed. When a blockchain is born, it could be tested with the agreed criteria to valid if the blockchain works fine as developers claim.

6.2 Big Data Analytics

Blockchain could also be combined with big data. Here we roughly categorized the combination into two types: data management and data analytics. As for data management, blockchain could be used to store important data as it is distributed and secure. Blockchain could also ensure the data is original. For example, suppose blockchain is used to store patients health information, the information cannot be tampered and it is hard to steal that private information. When it comes to data analytics, transactions on blockchain could be used for big data analytics. For example, we can extract user trading patterns to predict their potential partners' trading behaviours with the analysis.

6.3 Blockchain Applications

Currently, most blockchains are used in the financial domain but traditional industries could take blockchain into consideration and use the application of 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. For example, In Raleigh, a ridesharing startup offers an open marketplace where riders connect directly with drivers by leveraging blockchain technology.

7. CONCLUSION

The advent of this distributed, decentralized technology is taking information storage and security to new heights. With our generic front-end platform (restricted for now) which can take a variety of inputs (highly specific for now) for creating a transaction in a set of domains, the user can create a distributed, decentralized and an online transaction block, which is visible to everyone in the p2p network. Moreover, since it is highly secure, none of the agencies need to verify the data once the data is officially verified and added to a block. We believe that with comprehensive research on the consensus algorithm (that is used to accept the block into the transaction list to all the systems) and a robust verification system (for verifying documents), this technology can truly change the face of how sensitive information can be made fraud-proof.

8. ACKNOWLEDGEMENTS

We would like to extend special thanks to our Professor, Timothy Menzies and our mentor, Ken Tu for helping us to come up with this idea and giving us the opportunity to work on state of the art new technology.

9. REFERENCES

1. Blockchain Technology Beyond Bitcoin:
   http://scet.berkeley.edu/wp-content/uploads/BlockchainPaper.pdf
2. Blockchain technology inovations:
   http://ieeexplore.ieee.org/document/7998367/
3. https://www.quora.com/What-are-the-key-properties-of-the-Bitcoin-blockchain
4. An Overview of Blockchain Technology: Architecture, Consensus and Future Trends: http://ieeexplore.ieee.org/document/8029379/#full-text-section
5. The Digital Privacy Paradox: Small Money, Small Costs, Small Talk: https://papers.ssrn.com/sol3/papers.cfm