Blockchain based Video Conferencing System with Enhanced Data Integrity Protection Auditability

Nethmi Hettiarachchi
Department of Information Technology
British Collage of Applied Studies
Wallawatte, Sri Lanka

G.A.P.P.R. Pathiraja
Department of Information Technology
Sri Lanka International Buddhist Academy
Kandy, Sri Lanka

ABSTRACT
With the belief of every person should have access to a platform to share their voice, video conferencing is increasingly becoming the most genuine way to communicate with the trusted parties. But, due to the potential vulnerabilities of communication in-between entities exchanging information in a video conferencing systems may lead to catastrophes. In this research, a decentralized video conferencing systems is proposed to establish trusted communication in-between entities of the system and to enhance the integrity protection of stored information by utilizing the blockchain technology. As it is based on decentralized transaction ledger of blockchain, auditability, reliability of the exchanging information and the entities are enhanced. Each and every transaction of the system is traced and by performing intrusion detection, system is protected from malicious impersonators and intrusions.

Keywords
Block chain Decentralized Ledger

1. INTRODUCTION
Date integrity and auditability are critical security aspects in web based video conferencing applications. The goal of this research is to deliver a trusted protocol and platform for decentralized live video conferencing over the internet, to provide the open platform that gives users the easy ability to get their content and message out to the trusted party. Through coordination in a blockchain based protocol, solutions are possible that can result in a platform being: More secure, scalable, reliable and auditable, cheaper to the end user without any single points of failure. Through the combination of peer-to-peer network, and a soundly designed blockchain based crypto-economic protocol, this research aims to deliver the network that can accomplish all of the above for video conferencing.

A software system can be characterized into two main architectural approaches, i.e. centralized and distributed [1]. In centralized software system, the nodes are located around and connected with one central node of coordination. Distributed system, on the contrary, has several connected nodes without any central node of control. Fig. 1 illustrates the contrast of these two architectures. There are several benefits of a distributed system, i.e. having more computing power by combining the computing power of all connected nodes, an increased reliability due to the fact that it does not have a single of failure, and so forth. However, several drawbacks of a distributed system include communication overhead and security issues which is related to misuse network access by untrustworthy nodes.

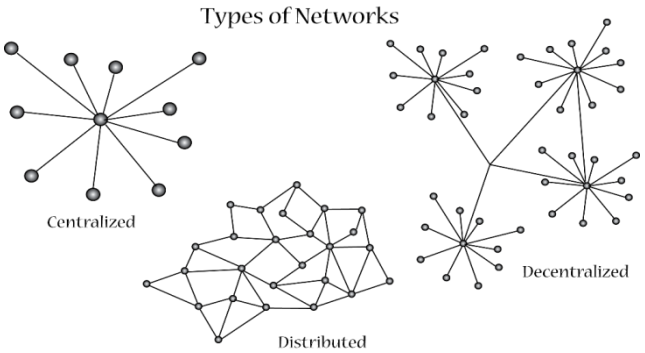


Figure 1: Centralized and Distributed Network Architectures

Blockchain is a type of distributed ledger (data structure) which contains information about transactions or events. It is replicated and shared among the participants in the network [2]. The size of chain unceasingly increases since blocks are added and chained to the previous block using a hash function (see Figure 2 for further illustration of the Bitcoin'sblockchain as an example).

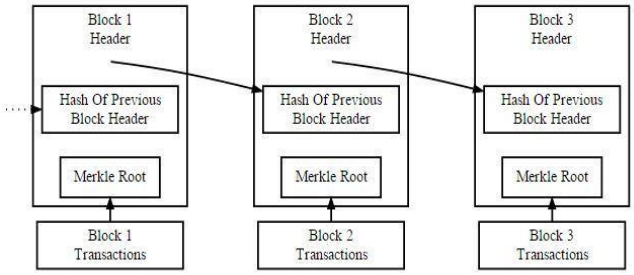


Figure 2: A chain of blocks in a blockchain

2. LITERATURE REVIEW

2.1.Livepeer Project
In "Livepeer Project" they discuss about a decentralized live streaming video protocol based on blockchain using "Etherum". The Livepeer project aims to deliver a live video streaming network protocol that is fully decentralized, highly scalable, crypto token incentivized, and results in a solution which is cheaper to an app developer or broadcaster than using traditional centralized live video solutions. Decentralized applications (DApps) to be built in the form of largely static or infrequently updated web or mobile content, but at the moment DApps still lack the ability to include streaming media and data in an open. and decentralized way . As the " Livepeer Project" claims, the goal of their project is to decentralize live video broadcast over the internet. Although it is a secure platform based on the

blochchain technology, "Livepeer Project" has not suggested any mechanisms to auditability and intrusion detection[3].

2.2.SteemQ

A decentralized platform for STEEM

SteemQ is a still on-going project aligned with "Steemit" decentralized social network which is based on blockchain technology. SteemQ is proposed to be a decentralized video platform for user- generated content based next-generation platforms on top of the new Blockchain and P2P technologies.

The blockchain of choice is STEEM. This allows the developers to build on top of the same technology that powers Steemit, as well as inherit the benefits of an existing community, currency and platform. All STEEM social network accounts are automatically SteemQ accounts and vice-versa. The system aims to empower its users to the maximum extent possible while remaining resilient. SteemQ say that it uses IPFS as a core building block of the content distribution system of their prototype. IPFS is a great tool that does a few things really well. It provides a robust layer for managing, transporting, referencing, deduplicating, versioning and ensuring the integrity of the content.

SteemQ suppose that they can secure the multihashes by storing them on the immutable blockchain (i.e. STEEM's posts become immutable after first reward payout. STEEM's transfers are much faster, being immutable and permanent within seconds, when the block is confirmed). This way they suggest that, they can simultaneously guarantee the ownership and integrity of the content [4].

2.3.Provr Chain

A Blockchain-based Data Provenance Architecture in Cloud Environment with Enhanced Privacy and Availability

In this paper, researchers have presented a concept called "ProvChain", a blockchain based data provenance architecture to provide assurance of data operations in a cloud storage application, while enhancing privacy and availability at the same time. ProvChain uses the construction of the merkle tree technology for the provenance of data. A list of blockchain transactions will be used to form a block and the block needs to be confirmed by a set of nodes in order to be included in the blockchain. An attempt to modify a provenance data record will require an adversary to locate the transaction and the block. Blockchain's underlying cryptographic theory will allow to modify a block record only if the adversary can present a longer chain of blocks than the rest of miners' blockchain, which is quite difficult to achieve [5]. As the "ProvChain" claims, the goal of their project is to provenance of data in the IoT based cloud environments. Although it is a secure platform based on the blochchain technology, "ProvChain" has not suggested any mechanisms to auditability and intrusion detection.

3. RESEARCH GAP

Based on the above literature, a clear research gap between the existing system upon the newly suggested system can be identified.

Noticing, there is a huge lack of decentralized video conferencing systems based on blockchain technology. There are very few on-going projects which have not demonstrated proper security and intrusion detection mechanisms, scalability and auditability such as SteemQ.

Although there are ongoing projects for live video streaming protocols like "Liverpool Project" based on blockchain technology, these applications haven't demonstrated any mechanisms to enhance the auditability and intrusion detection. Moreover, they haven't suggested any techniques to trace the information-exchange between entities and the validity of the transactions. In this kind of a background, there is a huge necessity of a secure, reliable decentralized video conferencing system. In this proposal a blockchain based video conferencing system with enhanced data integrity protection and auditing is presented to fill-up the research gaps of this domain.

4. METHODOLOGY

In this research, a blockchain based video conferencing system with enhanced data integrity protection and auditing is presented. To simulate the research idea, a web based e-health video conferencing application will be used. Figure 3 explains the basic architecture of the web based application.

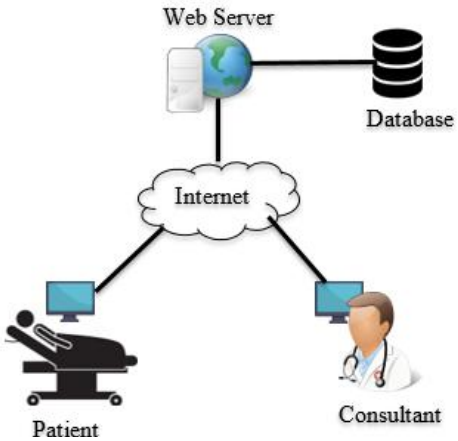


Figure 3: Basic system Diagram of the web application

Users

In the web application, there are two main user categories; patients and consultant doctors. Initially users (patients & consultant doctors) can register with the web application. Once they are registered with the web application, patients can look for the available consultant doctors. Then patients can book or reserve a particular doctor for a private session by paying the required amount of money by the application. When users register with the web application, the user is also registered and linked with the Application Control Module of the video conferencing system as well.

1. As shown in Figure 4, the system has four main components; User, Application Control Module, Blockchain Network and Storage and Validation Server. The main components of the system can be explained as follows;

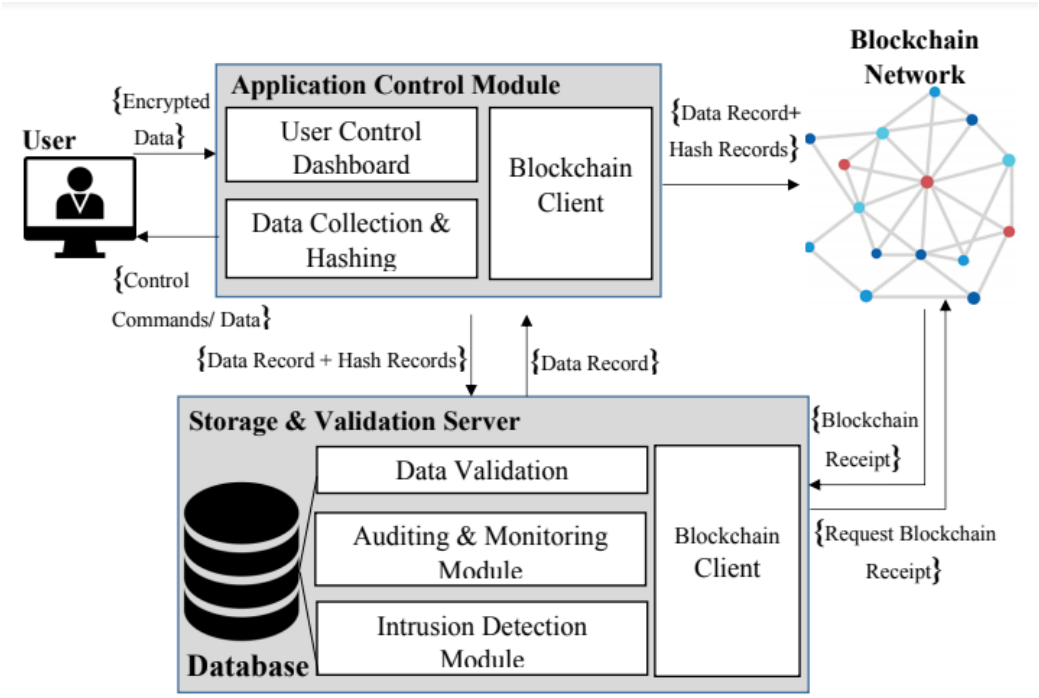


Figure 4: System Diagram of Blockchain Based Video Conferencing System

There are 2 main user categories as patients and the consultants, in this system. Both of the users can produce their own data such as videos, images or text files. As both the user categories are generating same type of data, they are equally treated in the system as "users".

The data records (videos/ images /text files) produced by the users should be sent to the Application Control Module in such a way that data is secure during the transmission.

4.1. Application Control Module

Data Collection and Hashing sub-component of the Application Control Module act as an intermediate to collect and receive data from the users. It is also responsible to hash the gathered data. Both hashed and the original data pieces are sent to the Storage & Validation Server and to the Blockchain Network.

Control Commands given away by the Application Control Module are also hashed and sent to Storage & Validation Server and to the Blockchain Network along with the original commands.

4.2. Blockchain Network

Blockchain is used to store data produced by the users & the commands generated by the Application Control Module. Hash records of the corresponding data pieces are also stored in the blockchain. Blockchain Network is also responsible to generate a blockchain receipt for the request made by the Storage & Validation Server for validation purposes.

4.3. Storage and Validation Server

Storage & Validation Server locally store data received from Application Control Module (Original data & corresponding hash records the local database as a key value store. For the validation purposes; Data Validation component periodically requests for blockchain receipt from the blockchain network. After the data validation the blockchain receipt along with the receipt request is stored in the local database.

All the database transactions and the data transmission between the Application Control Module and the blockchain are stored in the local database. As all the transmissions are traced, the video conferencing system itself is auditable. Security dashboard periodically and continuously observes the transactions & performs intrusion detection. Security dashboard will listen to operations and if any abnormal behavior of the component is detected, it alerts the control system and has authority to suspend or temporarily block the particular user from the video conferencing system.

Data records sent to the blockchain by the Control Module are treated as blockchain transactions. The records are hashed and they are used to construct a Merkle Tree. Construction of the merkle tree in explained in Figure 5 as follows.

After adding the new hash record to the Merkle Tree as a new Merkle Leaf, the value of the Merkle Root is calculated [6]. Calculated Merkle Root values then embedded in the new block generated by the blockchain. Embedding the merkle root into a block of a blockchain is explained in Figure 4.

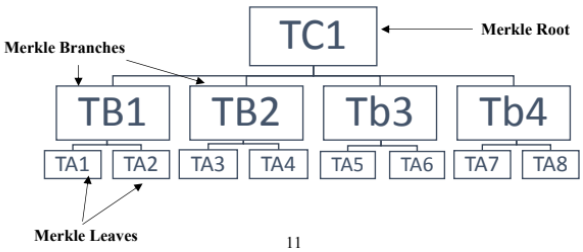
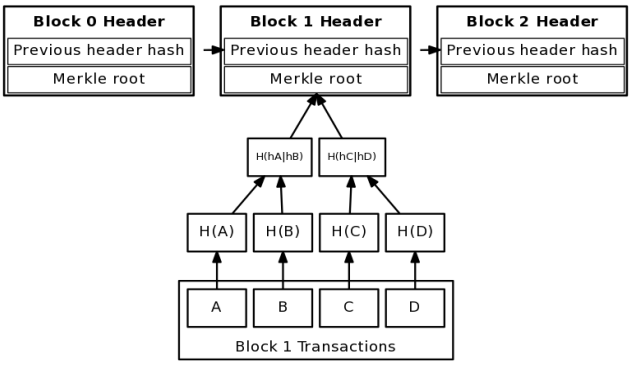


Figure 5: Construction of the Merkle Tree



Merkle tree connecting block transactions to block header merkle root

Figure 6: Embed the merkle root into a block of the blockchain

Each and every node of the blockchain should mine the block and validate to add the new block as a valid block to the chain.

4.4.Establishing the cryptographic keys

Any user who needs to communicate via the video conferencing system should be registered with the system (Application Control Module)

In the Storage and Validation Server, a key pair is required to encrypt the sensitive data (Session details). Required cryptographic keys are as follows.

I. User Registration key: K_User

With the registration of the user, the Key K_User will be created. Whenever a user performs any action within the system the data should be accessed via this K_User. Each and every time a new piece of data is being created, this K_User registration key is needed to store data.Viceversa, the registration key for the Application Control Module is K_ACM.

II. Data Encryption Key: K_DE.

After registration, the user generates an encryption key K_DE, for encrypting all the data. When a data entry is created, the user encrypts this piece of data entry, which limits the data access only to valid key holders. Each time there is a data entry created, the hashed data entry will be recorded on the blockchain.

III. Data Access Public/Private Key Pair (P_KDM, P_RDM)

For data access, a public/private key pair will be generated, denoted as (P_KDM, P_RDM). For some cases that the data access activity is to be recorded on the blockchain, the private key is used to generate a fingerprint from the operator to indicate the data origin, while the public key is used by others to verify the stated origin.

5. IMPLEMENTATION

In this video conferencing system, there are few steps where data is collected and transmitted among the main components of the system. User registration, session creation detail transmission, blockchain receipt generation, and validation of data..

A. User Registration

In the video conferencing system, the user needs to enroll as a node for storing data collected from a certain location. After registration, the data collection phase starts and a unique ID will be assigned to each user. Every data record will be associated with the user ID. The data type may contain some measurements, videos or images. For simplicity, we consider the data as an object and hash each data record for efficiency before uploading to the blockchain network. The original data is stored in a local database at the same time for future lookup.

B. Data and Command Transmission

Each time there is a data record collected from the user, the data entry can be constructed as a tuple {UserID, Time, Location, Data}. After the tuple is sent to the controller, the controller will forward the data to the blockchain network. At the same time, it will send back some commands based on the data and task. The commands will also be recorded on the blockchain, using the tuple {ControllerID, Time, Location, Command}.

C. Blockchain Receipt Generation

Once a collected data record from a user is uploaded to the blockchain network via the application control module, thisaction will be captured as a blockchain transaction. This provides the data management system ability for future validation, tracking and auditing. The record is hashed and eventually transformed into a Merkle tree node [8] using Tierion API [9]. The Merkle tree root node will be anchored in a blockchain transaction following the Chainpoint 2.0 protocol [10].

The use of the Merkle tree offers the scalability which satisfies the vast throughput from large numbers of users. A set of data records will be batched together as a transaction in the blockchain. A list of the transactions will be used to compose a new block, which will be confirmed by blockchain nodes. When the block is validated, it will be added to the existing blockchain, making it part of a tamper-resistant ledger. The blockchain receipt contains information of the blockchain transaction and the Merkle proof used to validate the transaction. A generated bockchain receipt is given in Figure 6.

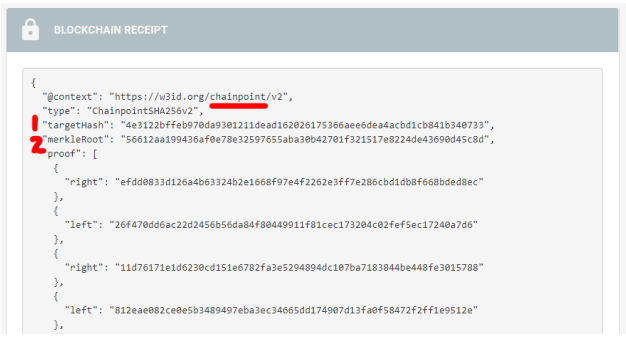


Figure 6: Generated blockchain receipt

D. Data Validation

Since each record will be stored in the cloud instantly, the data integrity can be verified at any time. By periodically requesting the blockchain network for a blockchain receipt, every record will be validated by comparing the calculated hash with the targetHash from the generated blockchain receipt [7]. If an inconsistency is detected, the record could be

suspected for compromise. A daemon process in the database server is configured to make a request via Tierion API [8], as a proof of integrity, using its URL[9].

The request header should include Content-Type: application/x-www-form-urlencoded or Content-Type: application/json to set the data format and the requests to the Data API is protected by HTTPS. The data records are validated with the input including targetHash, merkleRoot and the proof from the blockchain receipt. The most important step is to reconstruct the Merkle tree from the blockchain receipt to compute the Merkle root. Each data record is stored together with other records in the blockchain network as one transaction. The proof part of the receipt indicates the relationship between each record from the same transaction. For example, the left node means its record is collected earlier than the record anchoring in the right node. The transaction attribute height represents the block index, and we can find the exact block information in Block Explorer [10]. To validate the format and contents of a blockchain receipt, and to confirm that the Merkle root of one record is stored in the blockchain, the following URL provided by Tierion API is used.

6. SYSTEM IVALUATION

6.1.Analysis of Security

This video conferencing system usedblockchain technology concepts, which allows it's user for a secure communication and provide higher degree of privacy and auditability. The control system is an intermediate entity between the user and the storage server and responsible for forwarding hashed data towards the blockchain network. The commands along with the users data will be anchored to the blockchain network for integrity protection using blockchain receipts. By binding the userID and location information, the data source is trusted by an unalterable fingerprint. The database server hosts a database for real-time processing and provides persistent data availability.

Moreover, the server has the capability to integrate auditing module to inspect data and command records. The trusted data and command records contribute to the accountability of system components. By securing data process and distributing data process flow, high level of data assurance and resilience is preserved. This decentralized architecture in public blockchain network helps to provide robustness and tamper resistance for data assurance.

6.2.Performance Evaluation

To test the performance of this video conferencing system, a prototype has built to simulate the data collection and data transmission process. The collected data is uploaded from the control system to the blockchain network. The evaluation environment setup includes the server, data collection application, and a benchmarking tool.

A test plan has been built to measure the performance of the system. Test plan aims to simulate the action of uploading collected data to the blockchain using hashing algorithms. The simulation also uses random numbers to represent data content collected by theusers. The test plan contains one controller to generate HTTP POST request to the server. A different number of uers and different size of data are simulated to test the scalability of system.

Figure 7 shows the average response time of video conferencing system for data transmission application with a varying number of users, with the data size of 64 Byte.

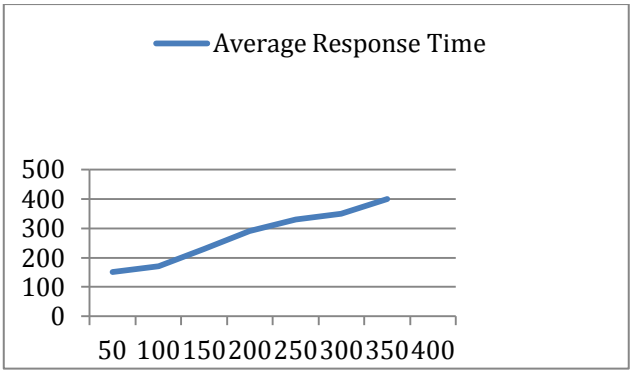


Figure 6: Average response time

7. CONCLUTION AND FUTURE WORK

The users always have the imitation of network dependency. Before submitting data to the real-time database, "Storage and Validation Server" stores data locally. But this may lead to create other security vulnerabilities. So, as future improvement, this loophole of using the local database should be eliminated and system should be supported with real-time data storage.

This system can protect the user's video session and the data that has been created by the user after the data generation. In future it will be worthwhile to secure the live streaming of the user's video session as it can contain highly sensitive content. This system is capable of providing reliability and accountability, as well as data assurance for real-time data collection and user control. In future, it is prominent if the system can be extended over a private andblockchain which will be more secure and protected. Dedicated private blockchain can increase the degree of offered security in a higher level.

8. REFERENCES

[1] B. A. Tama, B. J. Kweka, Y. Park and K.-H. Rhee, "A critical review of blockchain and its current applications," in 2017 International Conference on Electrical Engineering and Computer Science (ICECOS), 2017.

[2] C. P. Nabil El Ioini, "A Review of Distributed Ledger Technologies," in OTM Confederated International Conferences "On the Move to Meaningful Internet Systems", 2018.

[3] D. Petkanics, "Livepeer Proect overview," 2018.

[4] Furion, "SteemQ - A Decentralized Video Platform for STEEM," 2016.

[5] V. R. C. K. K. K. L. N. Sachin Shetty, "Data provenance assurance in the cloud using blockchain," in Proceedings Volume 10206, Disruptive Technologies in Sensors and Sensor Systems, 2017.

[6] U. C. &. S. Pandian, "HARE: A new Hash-based Authenticated Reliable and Efficient Modified Merkle Tree Data Structure to Ensure Integrity of Data in the Healthcare Systems," in Journal of Ambient Intelligence and Humanized Computing , 2021.

[7] S. B. H. Youssef, S. Rekhis and N. Boudriga, "A Blockchain based Secure IoT Solution for the Dam Surveillance," in 2019 IEEE Wireless Communications and Networking Conference (WCNC), 2019.

[8] "https://tierion.com/," 2018. [Online].

[9] "https://tierion.com/chainpoint," Tiarion API. [Online].

[10] X. Liang, S. Shetty, D. Tosh, C. Kamhoua, K. Kwiat and L. Njilla, "ProvChain: A Blockchain-Based Data Provenance Architecture in Cloud Environment with Enhanced Privacy and Availability," in 2017 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID), 2017.