Enhancing Data Security in Cloud using Blockchain

Dhananjay Yadav
Department of Information Technology
A.P. Shah Institute of Technology
Thane, India
dhananjayyadav98@gmail.com

Aditi Shinde
Department of Information Technology
A.P. Shah Institute of Technology
Thane, India
aditishinde228@gmail.com

Akash Nair
Department of Information Technology
A.P. Shah Institute of Technology
Thane, India
akashnair485@gmail.com

Yamini Patil
Department of Information Technology
A.P. Shah Institute of Technology
Thane, India
ympatil@apsit.edu.in

Sneha Kanchan
Department of Information Technology
A.P. Shah Institute of Technology
Thane, India
sskanchan@apsit.edu.in

Abstract:
Daily lots of data is exchanged and loaded on cloud into different sectors one of which is health sector. Data exchanged between the patient and doctors need to be secured to gain patients trust. Blockchain is a mechanism invented to secure data in more advance way. Blockchain store data into chunks that make it hard to decode, which will help provide extra layer of security. Hash chain is the most reliable part of blockchain that will help keep the data unreadable. This data can be secured by using a blockchain mechanism at the backend of any hospital website to store the reports of the patients, and maintain a two-way authentication for doctor's access to the reports. Using the concepts of dividing data into chunks and establishing an inter-link between each chunk is one of the aspects of blockchain which is implemented on the hospital generated data to inherit blockchain mechanism. In this paper we have discussed the benefits of using this mechanism to secure patients reports and how it increases trust on the stored data.

Keywords: **- Blockchain, Cloud, Healthcare, Security, AES**

I. Introduction

Despite using certain frameworks, we have implemented the idea through hard coding. Each report is divided into chunks and these chunks need to be secured in order to maintain integrity of data. Each of the chunks are encrypted through AES the key generation algorithm generates the public key (PK), master key (MK), secret key (SK) of user. There are K number of users in group sharing data. Master user is a owner of data. So, all the user can access and modify the shared data's in cloud. TPA performs data integrity auditing for modified data of the user. Key Generation as the part of set up algorithm generates public keys (PK), master keys (MK) of the system and secret key (SK) of users. In our design each user will have their own secret key for data modification Key generation is a technique which is used to store the data in a different methodology and mainly the public key algorithm known as RSA plays a vital role in key generation technique such as single shared key uses symmetric key algorithm through which data will be stored very secretly Since the public key algorithm uses two keys namely public and a private key and that public key is made as visible to one end user so that they can use that public key to encrypt the data and another end user can decrypt the data using their private key. In some conditions they keys have been generated using Random Number Generator technique, and it is very efficient that hackers cannot easily guess the keys and provide a strong security. the major focus in our paper is about the healthcare data. Hospitals generate a large amount of confidential data and especially every healthcare center needs to maintain HIPAA rules. Here we are using blockchain mechanism to the data that's been uploaded by the patients regarding their diseases and that data is then secured through blockchain mechanism. This security mechanism will be the backend processing of any hospital website, Recently, a few attempts started considering more realistic scenarios by allowing multiple cloud users to modify data with integrity assurance. Nevertheless, these attempts are still far from practical due to the tremendous computational cost on cloud users, especially when high error detection probability is required by the system. We used Amazon Cloud Storage service S3 for storing data divided into chunks in form of buckets.

II. Literature Survey

Blockchain originally refers to the increasing list of records or data in a chain like manner. The blockchain was found in

1992 as a technology to cease the tampering of timestamp on files. The latest technology blockchain can be traced back to 2008 when Satoshi Nakamoto (along with his colleagues) revised blockchain into a ledger like continuous transaction list which later led to the founding of the first crypto currency called bitcoin [1]. Bitcoin was the first currency to have blockchain backing it for its secure and tamper less transaction which led to crypto currency gaining a new advancement and boon to its use in today's date. Blockchain process involves breaking down transaction data into chunks or blocks of a sort which has three main parts that include data, previous hash and its own hash that has been generated by the hash of previous blocks hash and the blocks data. Any tampering to the block leads to the change in the hash values and which in turn leads the hash in the previous block to be untraceable. This makes the blockchain very useful to use in transaction and monetary related functions. The data in the block may include transaction details like sender, receiver, amount, ID etc. The blockchain technology seemed promising at first because of its high level of security that it provides but the cost and complexity installing it was very high. It wasn't possible to implement it everywhere because of its high perpetuity, complexity and network cost that comes with it. Hence the use of blockchain was limited to the crypto currency since it had a high risk while transactions because it's conducted through the network.

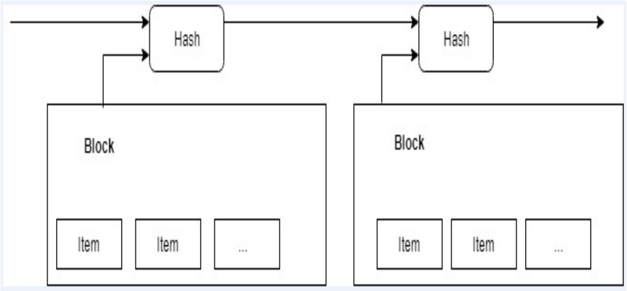

Figure No 1: Connection of Blocks

Amita Kashyap, G. Sravan Kumar, Sunita Jangir, Emmanuel S. Pilli, Preeti Mishra presented a model with IDS to maintain and check any kinds of irregularity in the system and network for detection of any inside or outsider attacks. Every year many companies report attempts of cyber-attacks on their resources. There are many kinds of attacks that cause users to worry about data confidentiality and integrity these include DDoS attacks, backdoor attacks, flooding attacks, phishing attacks, port scanning and various attacks on virtual machines and hypervisors. Any kind of intrusion attempts led to the cloud owner and admin getting the notification related to it. This made the system secure and led to possibility of immediate counter against the attacker. The one major flaw with including an IDS at hypervisor layer was that the IDS only notified and didn't counter the attack itself, it wasn't capable of blocking the attack on its own. IHIDS [2] contains mainly of these given components. The network packet capturing captures the live packets that are going through the virtual machine. The Wireshark packet sniffing tool is used for this purpose. The Wireshark is configured and embedded into the hypervisor layer to capture malicious data packets that are harmful to your system at the virtual bridge. The packet pre-processing includes holding off the .png files and extracting the header information from the data packets which is very useful for knowing the type of attack. After this, attack detection queries are fired which is used to identify the type of attack and generate a subsequent alert to the cloud admin and the cloud owner. System call tracing detects anomalies in the behavior by studying the logs. For this, an advance VMI tool named Drakvuf installed in the VMM layer. A modified version of hypervisor is installed which enables only certain permissible actions to be taken when dealing with the related resources like devices and memory. The trace pre- processing involves numbering each system call and storing them in a particular CSV file for further study of behavior and detection in the system. Finally, the alert generation and logging involve generation of logs based on anomalies in the normal procedures and also storing logs related to such in safe files. The relevant alert is also sent to the cloud admin and cloud owner.

In this referred paper the actual concept was to provide a consensus mechanism [3] to the data in which the publisher specifies that the mechanism was used to provide consensus mechanism to the data which is referred as contract data. A vast number of contractual documents, such as those relating to sales agreements, license papers, and other important documents are made all over the world. Current technology allows us to build record and manage these documents easily. However, the difficulty lies in protection of these documents for a major period of time as digital records make it easier to tamper the records and data. We believe that blockchain will help solve this issue. Because, blockchain itself is robust against attack, anyone can verify its records, and it's troublesome to vary its history. Hence the blockchain was implemented to have a contract transaction so secure that no data can be tampered or retrieved if attacker tries it so. The contract is divided in part of two where the mediator and contractors have a complete say in the part of contract. The contract is divided so that the data or address of the destination contractor, contract information, and the sender contract or information is all stored in hash values and sent through the network. The receiver contractor accepts and then the contract is sent back the first contractor to verify and which is then sent to the contract mediator who then confirms the contract which is then settled. This allows a complete trust and confidentiality between the contractors and mediator. It was one of the best ways of having an online transaction or contract which provided most security and anonymity to the contractor. The technology provided anonymity to the contractors which was itself a pro and a con. The anonymity implies that the data and information on the contractors remained unscathed and no information was leaked while the contract was in progress and no type of attack would breach the security of information, This was also a flaw in the system because the anonymity of the contractors means there is a possibility of fraudulent by the unknown contractors and it would be unable to trace them because the system itself maintains the anonymity of the contractors.

In the referred paper, it's mentioned that data breaches are one of the most common news cycle. Over the past 10 years, there have been so many breaches that the news get by and nothing can be done about it. Healthcare data is one of the few major ones to suffer from many frauds from simple identity theft to more complex schemes that involve forgery of fake hospital bills to selling of pharmaceuticals to black market. There are many kinds of fraudery possible if the sensitive medical health records are leaked. Unlike normal credit card and billing fraudery, the danger could be life threatening and may cause many problems to the victim in many ways possible. Since the early 2000's ERHs (Electronic health records) [4] have been adopted well over the world. According the aforementioned paper, the development of ERHs has been increased multiple times with well over 80% of the hospitals. Given the different types of data present for the medical records, it's not possible to come up with a single strong solution to secure the hospital data. It will involve a lot of cost and if the cost effective way were to be introduced, the security of the data could be a tradeoff. The paper mentions new use of technologies for security as the medical records and hospital sensitive data could be very important, hence new and improvement in securing the already available technologies are encouraged. Investments are made to make sure no security breaches occur that may lead to a lot of issues for the victims as well as organizations. Patients are also supposed to take active interest in the security of the hospital data and should make sure their data are being sent to the right hands through their means.

Collaborative work has been one of the few sorts after ways to work where multiple participants can work on the same project and have it available in a group. There's no better way for collaborative work but through data sharing and the means to this data sharing is provided by the cloud computing environment. The collaborative environment of cloud improves the efficiency of work in cooperative environments and has widespread potential applications. However collaborative work along with its perks has its own set of problems to deal with which is security of the data and how to efficiently share the outsourced data in a group manner. In the proposed architecture the authors have mentioned the use of key generation for a certain group of individuals which are needed to be formed into a collaborative group. The key is a common conference key [5] which all participants can use which sparingly increase as the number of participants increase in the conference. This allows the complexity of the key to be increased linearly in the conference as the participants increase and in fact also reduces the computational complexity that is needed for the key. It's an efficient way so the nullify different type of key attacks that are simulated on such groups.

III. EXISTING SYSTEM ARCHITECTURE

Existing use of blockchain is for electronic transactions of crypto currencies through public blockchain where these transactions are recorded as blocks of distributed network and a huge chain is maintained which requires great amount of computational power. Considering cloud, cloud has their own security rules and regulations such as IDS, Encryption methods, etc. But in the continuous increase the amount of data there lies certain loopholes that may tamper the data security.

IV. PROBLEM STATEMENT

In the current ratio of increase in breach by hackers on the data, Security breach on our data on cloud is something none of us would feel good about, and a proper effective solution with minimum requirement of computation is the current demand for every individual. Especially when it comes to important sectors like Healthcare wherein confidentiality is the root of working and large computational power is not an important aspect, maintaining the trust of patient considering HIPAA as well as the technology is the need of every healthcare organization.

V. PROPOSED SYSTEM ARCHITECTURE

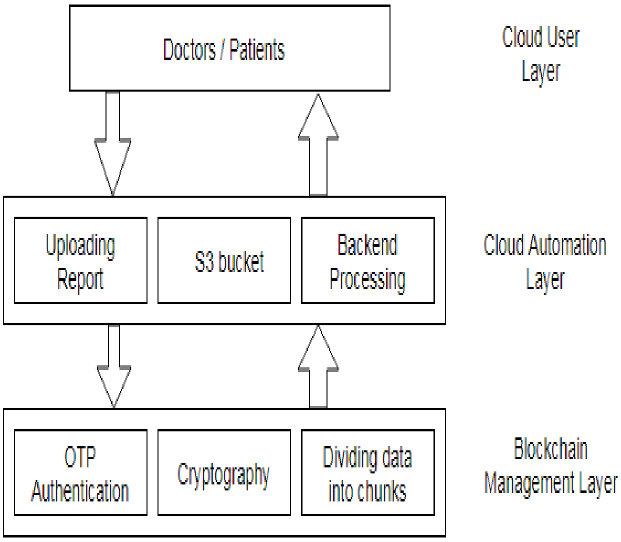

Figure 2: Flow of Proposed System

These are the primary objectives of our proposed system architecture:

- To divide the data into chunks.
- To maintain a proper encryption of the chunks.
- To have interlink between the chunks to form a chain of the data.
- Doctors must have a patient's side approval through OTP to access the reports.
- A proper GUI will be created which will help in the working for Doctors and Patients.
- Chunks will be stored into S3 buckets on AWS

Data Chunks: Dividing data into chunks will increase the data integrity, by making the data more complicated at the backend as a single report will be broken into smaller files each containing some part of the report.

Cryptography: AES Encryption fill help secure the data and maintain confidentiality, each chunk of the report will be encrypted through AES individually

Inter-linking: These chunks of report need to have some sort of connection to maintain authenticity of the data this can be achieved through linking each chunk with the next chunk following the concept of chaining of blockchain.
Consensus: Approval from the patient side should be maintained to achieve a trust between doctor and patient;

these can be done by maintaining OTP verification from patient's side for the doctors to download the report.

S3 buckets: Storage of these chunks of data needs to be in a secure environment and what's better than using a cloud environment, using S3 buckets to maintain the storage of file chunks. S3 makes to storage more reliable and durable.

The overall working of the proposed system will have an UI for any hospital through which the patients can upload their reports and doctors can view those reports. Doctors need to request the patients for downloading the report which requires an OTP that's been send on to the patient's email. As on for storage once the patient uploads the report the report is divided into chunks of file and these chunks are encrypted using AES, these chunks of data are distributed among 3 buckets that are formed using S3 an storage service by AWS. These buckets store the chunks of file randomly distributed among them and have an interlink to form a chain within the chunks.

VI. TECHNOLOGY STACK

Operating System	: Windows7 &Above
AWS	: Cloud Service platform.
AES	: Encrypting the data.
Navicat	: For visualization of the blocks and generated hash values.
Eclipse Luna for	: Integrated Development Environment java programming.
MySQL	: For Database storage.
JDK 1.6 &above	: For programming as it is easy to write and compile

Hardware Requirements:

System	: core i3 and Above.
Hard Disk	: 40 GB Minimum.
Ram	: 4 GB min.

VII. COMPARISON AND RESULTS

Our system provides an upper hand in the following ways:

- As we have developed an UI for hospital this gives the patients a flexibility of not carrying reports every time, they visit the hospital.
- Using cloud for storage purpose increases the efficiency and capacity of storage.
- Uploading reports as a softcopy lowers the paperwork of maintain hardcopies of each report.
- Normal data security encryption methods are replaced by advanced method such as dividing data into chunks and storing them in an encrypted form.

The current way of working in hospital adapts an traditional method of using hardcopy reports and mainly has no access of websites for the patients below table shows the regions our system covers compared to the existing system.

Sr No	Properties	Existing System	Proposed System
1.	Storing Reports in softcopy.	✗	✓
2.	UI for patients to interact with doctors.	✗	✓
3.	Using cloud service for storage.	✗	✓
4.	Storing Data in form of chunks.	✗	✓
5.	OTP Authentication from patient side to access reports.	✗	✓
6.	Encrypting reports	✗	✓
7.	Keep a note of prescribed medicines online	✗	✓

Table 1: Results drawn compared to Existing System

VIII.CONCLUSION

The main motive of our work is to create a trustworthy, efficient and real-time system for storing patients report in more efficient way, healthcare sectors rarely focus onto storing data into cloud and maintain the security of data and mostly prefer sticking to had copies of report, Our system not only lowered the task of maintaining hard copy records but also provided an security mechanism adapted through blockchain mechanism which will benefit a lot in maintain HIPAA Act more efficiently and help in achieving a growth in the way of working of health sectors.

ACKNOWLEDGEMENT

We would like to express our sincere gratitude to A.P shah institute of technology for giving us an opportunity to work under this design. We got an opportunity to learn about the working of blockchain and how security plays an important role in healthcare. We would like to thank our college faculties for arranging the necessary apparatus and guiding us throughout the process. We would like to thank our HOD Mr. Kiran Deshpande for assigning us with such an innovative idea to work upon. In the end we would also like to thank our parents for their emotional support and motivation.

References

[1] Nakamoto S. Bitcoin: A peer-to- peer electronic cash system[J]. Consulted, 2008.

[2] 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI) Amita Kashyap, G. Sravan Kumar, Sunita Jangir, Emmanuel S. Pilli, Preeti Mishra “IHIDS: Introspection- Based Hybrid Intrusion Detection System in Cloud Environment”.

[3] Watanabe, H., Fujimura, S., Nakadaira, A., Miyazaki, Y., Akutsu, A. & Kishigami, J. J. (2015). Blockchain contract: A complete consensus using blockchain. 2015 IEEE 4th Global Conference on Consumer Electronics (GCCEP).

[4] Matousek, K. (2008). Security and reliability considerations for distributed healthcare systems. 2008 42nd Annual IEEE International Carnahan Conference on Security Technology.

[5] Shen, J., Zhou, T., He, D., Zhang, Y., Sun, X., & Xiang, Y. (2018). Block Design-based Key Agreement for Group Data Sharing in Cloud Computing. IEEE Transactions on Dependable and Secure Computing, 1–1.

[6] Zhe, D., Qinghong, W ., Naizheng, S., & Yuhan Z. (2017). Study on Data Security Policy Based on Cloud Storage. 2017 IEEE 3rd International Conference on Big Data Security on Cloud (Big Data Security), IEEE International Conference on High Performance and Smart Computing, (HPSC) and IEEE International Conference on Intelligent Data and Security (IDS).

[7] Bharadwaj, D. R., Bhattacharya, A., & Chakkaravarthy, M. (2018). Cloud Threat Defense- A Threat Protection and Security Compliance Solution. 2018 IEEE International Conference on Cloud Computing in Emerging Markets.

[8] Suma, V. (2019). SECURITY AND PRIVACY MECHANISM USING BLOCKCHAIN. Journal of Ubiquitous Computing and Communication Technologies (UCCT), 1(01), 45-54

[9] Praveena, A., and S. Smys. "Ensuring data security in cloud based social networks." In 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA), vol. 2, pp. 289-295. IEEE, 2017.

[10] Homoliak, I., Venugopalan, S., Hum, Q., & Szalachowski, P. (2019). A Security Reference Architecture for Blockchains. 2019 IEEE International Conference on Blockchain (Blockchain).

[11] Killer, C., Rodrigues, B., & Stiller, B. (2019). Security Management and Visualization in a Blockchain-based Collaborative Defense. 2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC).

[12] Hendre, A., & Joshi, K. P. (2015). A Semantic Approach to Cloud Security and Compliance. 2015 IEEE 8th International Conference on Cloud Computing.

[13] Koo, J., Kim, Y.-G., & Lee, S.-H. (2019). Security Requirements for Cloud-based C4I Security Architecture. 2019 International Conference on Platform Technology and Service (PlatCon).