Blockchain Based Data Sharing Framework

Nitesh Singh[1], Jay Bothra[2], Sandeep Kothale[3]

[1]Nitesh Singh, Pune, Maharashtra, India
[2]Jay Bothra, Pune, Maharashtra, India
[3]Sandeep Kothale, Department of computer Engineering, SCOE, Maharashtra, India

---***---

Abstract - *Data is and will be a critical part of any modern application that we envisage. As we look into the future, the potential of machine learning & Artificial Intelligence will only make the volumes multiply. Data sharing comes with a wide range of challenges broadly categorized as: data format and meaning; legal obligations; privacy; data security; and concerns about unintended consequences of data sharing. It is not a trivial problem to solve as it does not pose only technical challenges but social, financial, ethical and regulatory as well. This creates the need to develop sharing frameworks which address technical challenges, embed regulatory frameworks, and anticipate and address concerns as to fairness and equity of outcomes in order to maintain trust of consumers and citizens. This paper proposes a framework that encompasses different ecosystems with respect to data sharing with blockchain technology as the backbone of this system. As blockchain inherently answers the major issues of trust, data accuracy and reliability, it goes on to provide a novel solution for data sharing.*

*Key Words***:** Blockchain, Public Key, Proof of Work, Distributed Ledger Technology,

1. INTRODUCTION

This document Domains like Medical/Healthcare, Banking, IoT are data-intensive domain where large amount is data (throughout this paper we are discussing about the structure data) is created, disseminated, stored, and accessed daily [6]. In last decade data from multiple source systems was collected, extracted, transformed and loaded into Datawarehouse (DWH). Later the data reconciliation is carried out between source system and DWH.

However after this data is created and when it needs to be shared this data over the network on internet we come across multiple challenges, data present with multiple parties would be different and not the same. This happens as most of the time data synchronization is not immediate and happens at end of day. This reconciliation process incurs cost. Also there is always a risk on tampering the data and in case of financial transaction double spending is always a concern.

1.1 Data Sharing on DLT

Here after studying the multiple papers (from as listed in references section) and various core components and functions of the Distributed ledger Technology (DLT) systems

which are characterized by secure, decentralized and distributed network. This paper proposes a framework that encompasses different ecosystems with respect to data sharing with blockchain technology as the backbone of this system. Blockchain inherently answers the major issues of trust, data accuracy and reliability; it goes on to provide a novel solution for data sharing. This further paper is organized as follows Section 2 describes data sharing challenges and work done, Section 3 describes how characteristics of Blockchain suits the framework with, the Proposed Datasharing Framework in Section 4. Finally Section 5 concludes with future Direction

2. DATA SHARING CHALLENGES

Traditional data sharing methods tend to be expensive for modern applications in terms of security, energy consumption and processing overhead. Most of the applications now a day's prefer to be on Cloud mainly due to convenience of high availability and saving in cost. There are, however, limitations to using conventional cryptographic primitives and access control models to address security and privacy concerns in an increasingly cloud-based environment. [6]

Also many of the state-of-the art frameworks are highly centralized and are thus not necessarily well-suited for modern applications data sharing due to the difficulty of scale, Security, no consensus mechanism, and single point of failure and user privacy and consent. Consequently, these modern Applications demands a consensus based data sharing, tamper proof, time stamp based data saving and distributed security and privacy safeguard. Next from literature survey we have listed the work done on Data sharing framework in different domain.

[1] Delivers a usable blockchain based model for a collection of researchers' data, providing accountability of access, maintaining the complete and updated information, and a verifiable record of the provenance, including all accesses/sharing/usages of the data.

[2] Describe the design and implementation of a smart contract for consent-driven and double-blind data sharing on the Hyperledger Fabric blockchain platform

[6] Study the potential to use the Blockchain technology to protect healthcare data hosted within the cloud. Also describe the practical challenges of such a proposition and further research that is required.

[8] Offers an overview of DLT and examines its potential for resolving the current problems of the KYC process and to achieve optimized KYC costs. Demonstrated how applied design science research to solve the problem. Also suggest possible implementation of solution.

[9] Propose a decentralized storage with blockchain in a common cloud based platform that is applicable to every type of Meta Products and ensures trust and privacy for sharing user data across applications developed by different manufacturers.

3. BLOCKCHAIN

Broadly speaking, blockchain is a technology able to build an open and distributed online database, which consists a list of data structures (also known as blocks) that are linked with each other (i.e. a block points to the following one, hence the name blockchain). These blocks are distributed among multiple nodes of an infrastructure, and are not centrally stored. Each block contains a timestamp of its production, the hash of the previous block and the transaction data. [6]

Blockchain has been proposed as a data structure used to create a public distributed digital transaction ledger which, instead of resting with a single provider, is shared among a distributed network of computers. Blockchain itself is similar to cloud computing where all the nodes are connected through the internet and all of them do the similar task of computing with which they will get rewarded. [9]

To summarize the blockchain provides the characteristics like Immutable Distributes and synchronized ledger across network, consensus mechanism to unanimously agreed between the parties/participants in network, uniqueness of data as it cannot be copied or replicated, data is time stamped with fully connected chronological blocks with audit trail, data is digitally sealed with cryptography and data is secure and tamper proof

4. PROPOSED BLOCK CHAIN BASED DATA SHARING FRAMEWORK

We propose a data sharing framework based on block chain technology. There are different parties involved here like Issuer, verifier. These are the users who may be either individuals or organizations who would add, amend or retrieve this data from the system. Intention of user is to upload the latest documents or artifacts to Block chain based system.

Here the data is referred to as records or documents to be updated which is the new or existing data, stored on cloud storage. This data can be requested by the user for analyzing and verifying purpose. System accepts this data and generates the hash for the document. This hash along with document is stored on cloud data storage. Post this document details along with hash is updated in the block and this block is delivered to all authorized peers only. This authorization can be updated by the authorization update module in the system. Once the block is verified and approved by the application the data is inserted in the chain and linked with previous block.
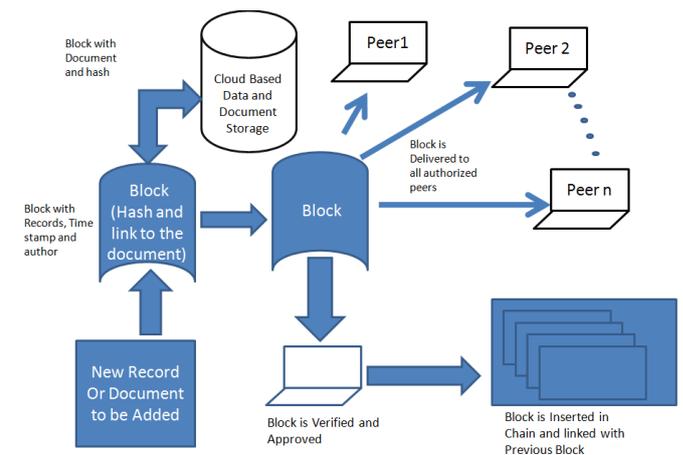

Fig -1: Proposed BC Based Data Sharing Framework

In the system there would be different roles depending on the use case. Let's say for data sharing in VISA application for particular country one role would be for country's embassy office, another role for the visa applicant and so on as per the requirement in usecase. Next comes the data format, here we suggest to use structured data which could be used with any relation database.

The application could be inter portable. All access right given in the system would be managed from the application based on the roles and these roles can be assigned to the user.

Most importantly blockchain helps us to manage the CI (Confidentiality and Integrity) aspect of security. For Availability the redundancy can be used at application architecture level and have not been explained in this paper.

Next steps are implementation of this framework and study it's the metrics of framework from further optimization and scalability aspect

5. CONCLUSION

IN this paper we proposed the block chain based data sharing framework which is very generic and can be applied to any domain where data sharing of sensitive data between multiple parties is a challenge. With this framework data is shared in secured manner and also the data privacy is

protected. The communication and authentication protocols need be further investigated and extend this research work with further exploration. In our future work we will continue with the implementation of the system based on this framework and conduct a study to get better optimization and empirical data which could be used for further exploration and studies

ACKNOWLEDGEMENT

REFERENCES

[1] Ajay Kumar Shrestha(&) and Julita Vassileva "Blockchain-Based Research Data Sharing Framework for Incentivizing the Data Owners" Springer International Publishing AG, part of Springer Nature 2018 S. Chen et al. (Eds.): ICBC 2018, LNCS 10974, pp. 259–266, 2018.

[2] Kumar Bhaskaran, Peter Ilfrich, Dain Liffman, Christian Vecchiola, Praveen Jayachandran, Apurva Kumar, Fabian Lim, Karthik Nandakumar, Zhengquan Qin, Venkatraman Ramakrishna, Ernie GS Teo, Chun Hui Suen± "Double-Blind Consent-Driven Data Sharing on Blockchain" 2018 IEEE International Conference on Cloud Engineering

[3] Z. Zheng, S. Xie, H. Dai, X. Chen and H. Wang, "An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends" 2017 IEEE 6th International Congress on Big Data

[4] KONSTANTINOS CHRISTIDIS, (Graduate Student Member, IEEE), AND MICHAEL DEVETSIKIOTIS, (Fellow, IEEE) "Blockchains and Smart Contracts for the Internet of Things" IEEE Access, Digital Object Identifier 0.1109/ACCESS.2016.2566339

[5] Ashiq Anjum, Manu Sporny, Alan Sill "Blockchain Standards for Compliance and Trust" IEEE CLOUD COMPUTING PUBLISHED BY THE IEEE COMPUTER SOCIET Y

[6] Christian Esposito Christian Esposito University of Salerno Alfredo De Santis University of Salerno Genny Tortora University of Salerno Henry Chang University of Hong Kong Kim-Kwang Raymond Choo University of Texas at San Antonio "Blockchain: A Panacea for Healthcare Cloud-Based Data Security and Privacy?" IEEE Cloud Computing Copublished by the IEEE CS and IEEE ComSoc January/February 2018

[7] Bin Yu Monash University, CSIRO Data61 Jarod Wright, Surya Nepal, Liming Zhu CSIRO Data61 Joseph Liu Monash University Rajiv Ranjan Newcastle University "TrustChain: Establishing Trust in the IoT-based Applications Ecosystem Using Blockchain " IEEE Cloud Computing Copublished by the IEEE CS and IEEE ComSoc July/August 2018

[8] Jose´ Parra Moyano and Omri Ross "KYC Optimization using Distributed Ledger Technology" Springer. Bus Inf Syst Eng 59(6):411–423 (2017)

[9] Ajay Kumar Shrestha(&) and Julita Vassileva "Towards decentralized data storage in general cloud platform for meta-products" BDAW'16, November 10-11, 2016, Blagoevgrad, Bulgaria © 2016 ACM. ISBN 978-1-4503-4779-2/16/11

[10] Wazen M. Shbair _, Mathis Steichen _, J´erˆome Franc¸ois y, Radu State _ "Blockchain Orchestration and Experimentation Framework: A Case Study of KYC" https://github.com/wshbair/Blockchain-Orechtration-Framework

[11] Satoshi Nakamoto "Bitcoin: A Peer-to-Peer Electronic Cash System" www.bitcoin.org

[12] Richard Gendal Brown, James Carlyle, Ian Grigg, Mike Hearn "Corda: An Introduction" http://r3cev.com/s/corda-introductory-whitepaper-final.pdf, 2016.

[13] Vimalkumar Pachaiyappan R. Kasturi "Block Chain Technology (DLT Technique) for KYC in FinTech Domain: A Survey" International Journal of Pure and Applied Mathematics Volume 119 No. 10 2108, 259-265