KYC As A Service (KASE) - A Blockchain Approach

Dhiren Patel, Hrishikesh Suslade, Jayant Rane, Pratik Prabhu, Sanjeet Saluja, Yann Busnel

KYC As A Service (KASE) - A Blockchain Approach

Dhiren Patel[1], Hrishikesh Suslade[1], Jayant Rane[1], Pratik Prabhu[1], Sanjeet Saluja[1], Yann Busnel[2]

[1] VJTI Mumbai, India
[2] IMT Atlantique Rennes, France

Abstract: KYC or Know-Your-Customer is an integral part of the onboarding process of a customer for a company. This process requires independent and tedious verification of a customer's identity documents by the businesses leading to wastage of resources. In this paper, we propose a solution where the submission and verification of a customer is done only once, and the results are shared with the businesses which require the information. The proposed system uses blockchain to record and manage the KYC requests and ensure transparency. The KYC data is verified using machine learning processes to ensure further efficiency in the process by reducing a significant amount of time spent on verifying the customers.

Keywords: Blockchain, Ethereum, IPFS, KYC system, Cryptographic Hash.

1 Introduction and Background

1.1 Know-Your-Customer (KYC)

Know-Your-Customer (KYC) refers to the steps taken by a business to establish customer identity, understand the nature of a customer's activities and to assess risks (if any) involved with the customer. It is a legal requirement for the financial institutions for on-boarding a customer. KYC requires the submission of the identity documents by the customer to the businesses or organizations on which they wish to onboard. Individual verification of the documents is done and thus establishing the identity of the customer independently.

Know-Your-Customer (KYC) is used for customer management and identity verification. This document is submitted by the customer to an organization for authentication and verification purposes.

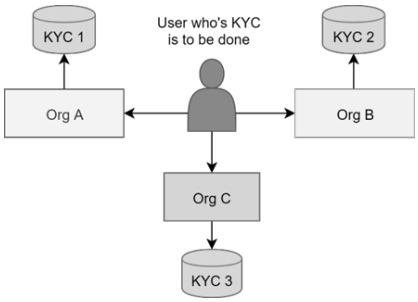

Fig. 1. Current KYC Implementation

Currently, KYC is done individually by every business and the same data is provided by the users to multiple businesses and independently verified by each of them. It would be far more efficient if the KYC could be shared securely amongst the organizations and hence give a better Quality of Experience (QoE) to the customer. However, due to the lack of trust between organizations this data is not shared between them, hence requiring a solution that can guarantee trust and reliability.

1.2 Blockchain

Blockchain is a distributed ledger technology that is used to ensure trust and reliability since the data and transactions are committed into the blockchain only after a consensus is reached amongst the participants. There are various consensus mechanisms that have been implemented to ensure a reliable distributed consensus.

Interplanetary File System is a distributed file system that stores files in a decentralized, distributed manner [1]. Blockchain has many applications and some of the use case are

1. Industry and IOT: Major use case under this topic include supply chain management, healthcare – patient data management, smart power grids [2], Agriculture – Agriculture food traceability and manufacturing industries [3].
2. Others: Creation of middleman free services such as blockchain based auction mechanism [4] and blockchain based death wills.

Until now, no centralized KYC verification system exists due to the lack of trust between institutions requiring individual and separate KYC processes and systems followed in each of them internally. Therefore, using a decentralized open technology such as blockchain would help ensure trust and integrity [5] from the ground- up and help in the open acceptance of this system.

Rest of the paper is organized as follows: Section 2 provides design rationale for KASE with the detailed architecture of KASE is discussed in Section 3. Section 4 gives prototype implementation of KASE using Solidity smart contracts with in-progress validation. Paper is finally concluded in Section 5 with references at the end.

2 KASE Design Proposal

The system proposed in this paper performs KYC As a Service — a service which acts as a one-stop solution for all of a customers' and businesses KYC needs. The customer provides the data to the service, where the service verifies the data using Machine Learning techniques and stores its encrypted format on a distributed file system and stores every transaction of the KYC data on a blockchain. A blockchain is a decentralized data structure in which transactions are conducted only after certain consensus is reached through consensus mechanisms. The proposed system uses an Ethereum Blockchain [6] with a Proof-Of-Work consensus mechanism. This mechanism allows the blockchain to enforce "smart contracts" such that the transactions are only committed to the blockchain after certain conditions are satisfied.

The system initially asks the user to register on the service and provide his information, including identity proofs to the service voluntarily. The next time the customer wants to get on boarded onto a business he/she uses the service for the KYC process. This information is stored encrypted by the user's secret key on the distributed file system and the transaction is stored on the blockchain to ensure transparency.

If a customer wants to onboard to a business, he/she can register to the business using the service and provide basic details which would be given to the business and verified by the service.

The service first asks the customer for confirming and validating the KYC request in accordance with GDPR and then after receiving the permission verifies the customer's identity to the business. The request transaction is also pushed onto the blockchain to ensure transparency of the data flow and credibility of the transfer.

The service also provides businesses the feature of verification of any KYC documents they may request based on their internal policies and uses Machine Learning approaches to verify those documents and to confirm the identity of the individual.

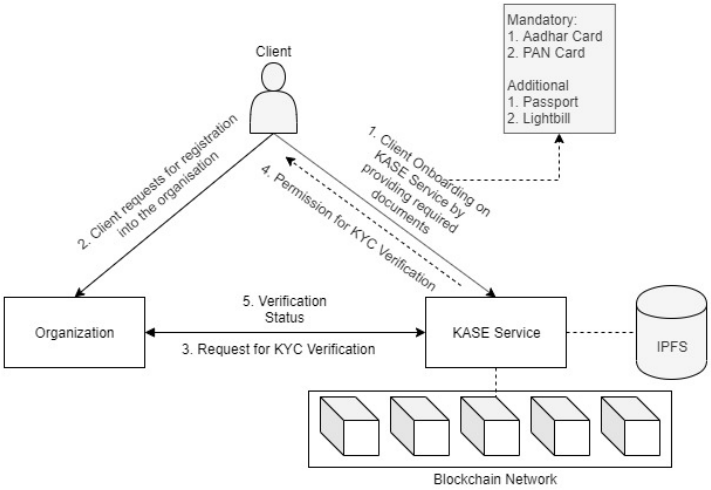

Fig. 2. High Level Architecture

The Architecture proposed in this paper (Fig. 2) for the Blockchain-Based KYC system uses a decentralized database (IPFS). We also use Machine Learning based Image Processing and Data Extraction for Legacy KYC processes. Resubmission

3 KASE - Detailed Architecture

At the time of on-boarding onto the system (KASE), the user will have to provide his identity proofs he/she has. The user will have to fill in all the details manually also, which will be converted to a JSON object. The data entered by the user and data extracted from the documents uploaded will cross checked for any irregularities using machine learning, and an extra layer of verification can be added by comparing the images of the user on various IDs and their image taken digitally.

Once all the checks are complete and all the data is verified, a public-private key pair is generated on the user's system. For security purposes the private key will be a key-file which could be stored on an isolated storage device such as a USB drive or a Gemalto Token.

The data stored in JSON object will be stringified so that it can be stored in IPFS along with the various ID documents, which are encrypted using the user's public key and stored on IPFS. All the documents that will be uploaded will have a different hash. The JSON file will also have a different hash.

All these set of hashes will along with the username will be stored in Ethereum blockchain as a “KYC on- boarding request”. The Ethereum wallet address generated will be of 42 characters which is impossible to remember. A mapping functionality provided by Solidity can be used to map username with the wallet address. At the time

of data retrieval, the user has to only provide his unique username to access his/her details. From username, the wallet address can be accessed and through that one can get their stored data.

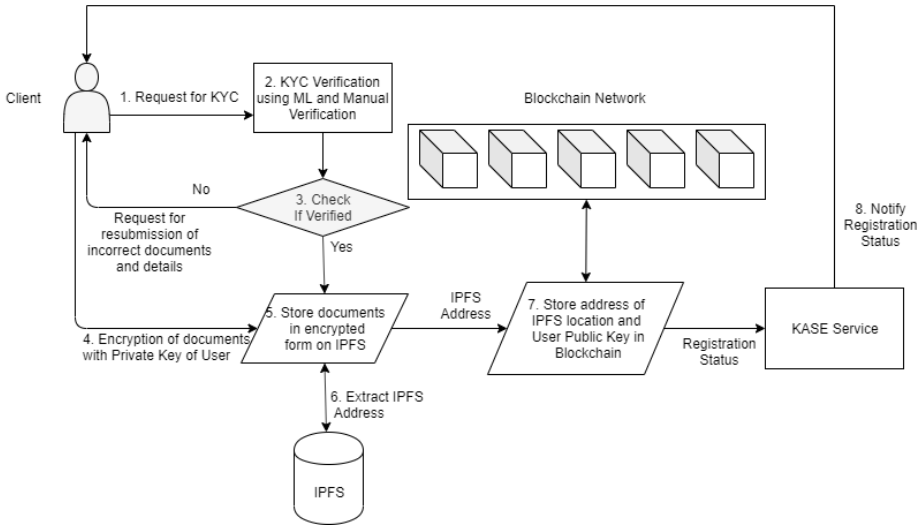

Fig. 3. User on-boarding on KASE

If a business wants to do the KYC of a customer, it can use the proposed service in two ways:

1. It can either request directly for verification
2. It can request the customer for documents and get those verified with the service.

When the business wants to KYC a customer, the business sends a request to the customer to allow the KYC to be processed by the business. KASE sends a notification to the user that the business is requesting KYC and the customer has to authenticate and allow the service to use the user's information to verify it to business. The system gets the address of the user's encrypted information from the blockchain and uses it to verify the customer details provided. After completing the request, the system pushes a "request transaction" to the blockchain.

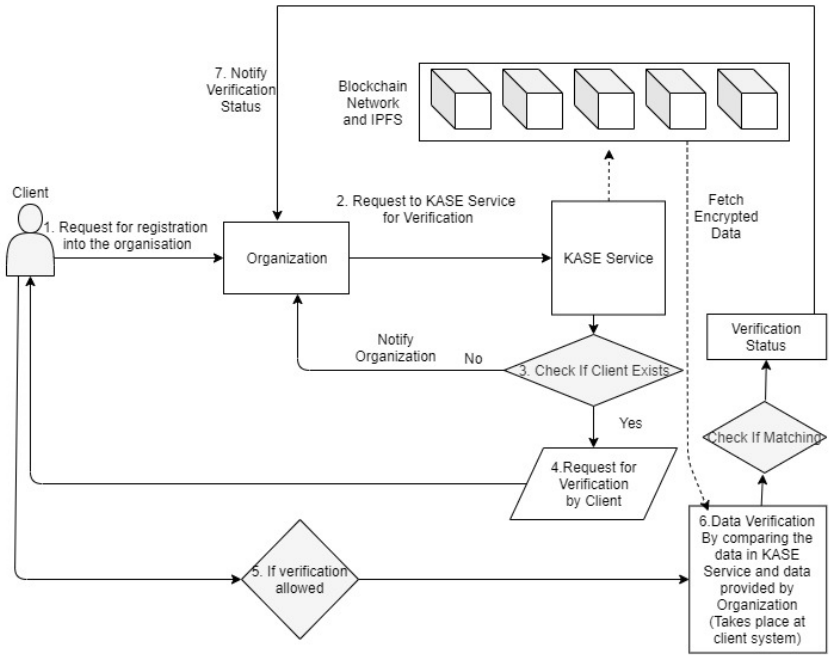

Fig. 4. KYC Process

In case of an event of a change in a user's KYC documents or details, the user has to provide the changed identity proofs to the system where the system verifies and ensures their credibility. The system finds the location of the previous documents using blockchain and pushes the new documents into that location. After completion of this request, this "update documents request" is pushed on the blockchain.

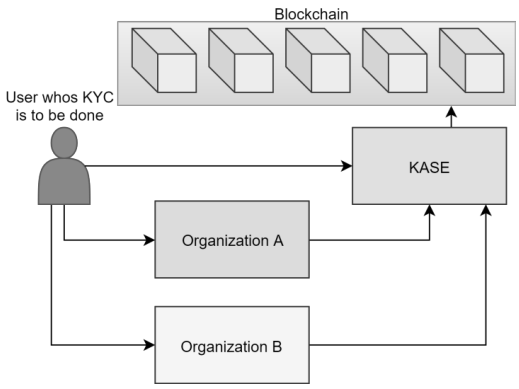

Fig. 5. KYC as a Service

If the business requests documents from the customer, KASE sends a request to the customer to authenticate and allow the request, and the service then retrieves the location of the encrypted documents provided at the time of on-boarding and uses Machine Learning to verify the currently submitted documents to ensure credibility. After completing the request, the system pushes a "request transaction" to the blockchain.

4 Prototype Implementation and Validation

Smart Contracts are written in Solidity which will be deployed on Ethereum Blockchain. The solidity version used is 0.4.21. There are two main Smart Contracts used.

4.1 Customer Contract

This contract is invoked when the customer is successfully verified and can be onboarded on the KASE system. It has a total of 6 functions.

1. To add data to blockchain
2. To get customer name
3. To get customer data link (IPFS)
4. To get customer Aadhaar Image link (IPFS)
5. To get customer PAN Card Image link (IPFS)
6. To get customer Passport Image link (IPFS)

4.2 Organization Contract

This contract handles the functionalities of onboarding or adding an organization to the KASE system. This contract has three functions.

1. To add an Organization
2. To get Organization name
3. To get Organization Details Link

4.3 Validation

The proposed solution leverages the core features of blockchain to ensure trust between the users, businesses and services to ensure the usage of the service. The service provides the following key advantages:

1. Cost effective solution for a business's KYC needs,
2. Saves valuable time of the business to ensure a customer's reliability.
3. No single point of failure of the system due to the usage of inherently decentralized components. It also achieves immutability of user data.
4. Ensuring openness, trust and reliability through blockchain.

5. Follows GDPR guidelines [13] by ensuring that the user data is not used without permission of the user.

5 Conclusion and Future Scope

Blockchain is one of the latest technologies in the field of cybersecurity and ensures trust in trustless environments. The proposed Blockchain-Based KYC system that uses a decentralized database (IPFS), Machine Learning based Image Processing and Data Extraction for Legacy KYC processes. Through blockchain, KASE ensures that the parties using the service can trust the service and its reliability, and will use it over other solutions. The solution further uses a decentralized file store to ensure complete decentralization of data and reduce any single points of failure. Our prototype implementation through Solidity smart contracts gives encouraging results.
KASE Service can be used as a one stop solution of all KYC needs. By leveraging the power of ML, AI and explainable AI we can make the system free of manual verification.

References

1. "IPFS-Content Addressed, Versioned, P2P File System." 14 Jul. 2014, https://arxiv.org/abs/1407.3561.
2. T. Alladi, V. Chamola, J. J. Rodrigues, and S. A. Kozlov, “Blockchain in smart grids: A review on different use cases,” Sensors, MDPI, vol. 19, no. 22, p. 4862, 2019.
3. T. Alladi, V. Chamola, R. Parizi, and K. K. R. Choo, “Blockchain applications for industry 4.0 and industrial iot: A review,” IEEE Access, Nov. 2019.
4. V. Hassija, G. Bansal, V. Chamola, V. Saxena, B. Sikdar, "BlockCom: A Blockchain based Commerce Model for Smart Communities using Auction Mechanism", IEEE ICC, Shanghai, China, May 2019.
5. "Blockchain And The Future of the Internet: A Comprehensive...." 23 Feb. 2019, https://arxiv.org/abs/1904.00733.
6. Wood, Gavin. "Ethereum: A secure decentralised generalised transaction ledger." Ethereum project yellow paper 151.2014 (2014): 1-32.
7. Nakamoto, Satoshi. (2009). Bitcoin: A Peer-to-Peer Electronic Cash System.
8. Vitalik Buterin, “The meaning of decentralization”, February 2017, Reterived from https://medium.com/@VitalikButerin/the-meaning-of-decentralization-a0c92b76a274
9. Sinha, Prince & Kaul, Ayush, Decentralized KYC System (2018) https://www.irjet.net/archives/V5/i8/IRJET-V5I8204.pdf
10. Parra Moyano, J. & Ross, O. Bus Inf Syst Eng (2017) 59: 411. https://doi.org/10.1007/s12599-017-0504-2
11. TCS Whitepaper - Reimagining KYC with Blockchain Technology, last accessed 2019/11/08.
12. Ali, A., Latif, S., Qadir, J., Kanhere, S., Singh, J., & Crowcroft, J. (2019). Blockchain And The Future of the Internet: A Comprehensive Review. arXiv preprint arXiv:1904.00733.
13. GDPR European Union Guidelines https://gdpr-info.eu/, last accessed 2019/11/08.