KYC Optimization Using Distributed Ledger Technology

José Parra-Moyano and Omri Ross*

August 4, 2017

Abstract

The know-your-customer (KYC) due diligence process is outdated and generates costs of up to USD 500 million per year per bank. We propose a new system, based on distributed ledger technology (DLT), that reduces the costs of the core KYC verification process for financial institutions and improves the customer experience. In the proposed system, the core KYC verification process is only conducted once for each customer, regardless of the number of financial institutions with which that customer intends to work. Thanks to DLT, the result of the core KYC verification can be securely shared by customers with all the financial institutions that they intend to work with. This system allows for efficiency gains, cost reduction, improved customer experience, and increased transparency throughout the process of onboarding a customer.

Keywords: Blockchain, Know your customer, Banking

1 Introduction

The increased regulatory cost incurred due to the know-your-customer (KYC) verification process in banking is one of the largest challenges that the banking sector is currently experiencing. The yearly direct costs that financial

*José Parra-Moyano is at University of Zurich, Moussonstrasse 15, 8044, Zurich, Switzerland, e-mail:jose.parramoyano@business.uzh.ch; Omri Ross is at University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark, e-mail:o.ross@statslab.cam.ac.uk; We would like to thank Ingrid Bauer, Roman Beck, Roman Dahl, Mads Garff, Mads Gram, Jakob Hedegaard, Fritz Henglein, Klaus Langenheldt, Catherine A. Schlicht-Bang, Karl Schmedders, Thorkil Vaerge, Johannes Wotzka, and seminar participants at the 2016 Blockchain Summer School, the Nordic Blockchain Summit, and Copenhagen University Hiperfit Seminar, as well as Ante Plazibat for excellent research assistance, comments, and help with the development of our ideas.

institutions need to cover in order to meet their obligations in terms of KYC are estimated, in a recent survey by Thompson Reuters (2016), to average USD 60 million[1]. This cost can be further augmented by the fines levied on financial institutions due to their misconduct with regard to anti-money-laundering (AML) and KYC regulations. According to the head of Strategy and Risk at the Hong Kong Securities and Futures Commission, *"KYC and AML stand out* [for a bank to] *as a pretty significant inefficiency and problem case [...] tallying up the fines [for a bank to] 10 billion or more US dollars"* (Benedict N. Nolens, at the MIT Technology Review Emtech conference, 2016). And the sources of additional costs do not stop here, as financial institutions are not allowed to conduct any business with corporate entities that have not yet completed the full KYC process. Since that process is long, and tends to lengthen with the size of the corporate entity concerned, the starting point of a given business relationship between a customer and a financial institution is usually delayed, which represents opportunity costs for both parties. Indeed, corporations need to verify all their subsidiaries before being granted KYC verification, and this is a laborious task for them. Therefore, it comes as little surprise that the abovementioned survey indicates that 89 percent of customers do not have a good KYC experience.

The aim of this paper is to propose a new approach to the KYC verification process. We introduce a system, based on DLT, that proposes a solution to the increased costs of the KYC process and the lack of customer satisfaction. The key reason for using DLT is that it allows us to observe the KYC cost structure at an aggregate level for all the financial institutions operating in a jurisdiction and to tackle the inefficiencies that emerge from the duplicated conduct of similar tasks by all participating institutions (i.e., DLT allows us to render the execution of duplicated tasks completely unnecessary, and this delivers far greater cost savings than would any effort to merely make these duplicated tasks more cost efficient). Specifically, DLT enables the creation of a chronological, decentralized, interbank ledger in which financial institutions that need to conduct the same KYC verification tasks for that customer can verify the result of the process that has already been conducted for that customer, thus avoiding conducting duplicated KYC verification tasks. Moreover, the use of DLT allows the cost of the KYC process to be shared proportionally among the financial institu-

[1] Average cost among the nearly 800 institutions that responded to an authoritative Thomson Reuters survey on the impact of global changes in Know Your Customer (KYC) regulation.

tions that work with a specific customer. In particular, the system allows customers to carry out the full KYC process with only one financial institution, and later on to share the result of that KYC process with any other financial institution that they intend to work with. The DLT acts as a "single point of truth", understood as the only source of information, accepted by any involved party should conflict occur.

The main improvement of the proposed system over the current system is that the KYC process only needs to be carried out once by each customer, rather than once by each institution working with that customer. This reduces the aggregated cost of the KYC process as a whole in a jurisdiction without compromising the security of the system, respects the privacy of the participants, and increases transparency in case of a conflict. Additionally, the use of the public key of a customer as a reference point for an immutable exchange of information across participating institutions serves as a basis for interbank collaboration. The use of DLT reduces the aggregate cost of KYC and this is the main conceptual contribution of this paper. In Section 2 we explain the KYC process, and relate it to work that has already been carried out with regard to optimizing KYC costs. Section 3 offers an overview of DLT and examines its potential for resolving the current problems of the KYC process. In Section 4 we show how we have applied design science research to solve the problem at hand. In Section 5 we describe and analyze the prototype solution and the economic mechanisms that need to be put in place in order to ensure a well functioning system. In Section 6 we discuss three possible implementations of this solution. Section 7 concludes.

2 The Current KYC Process

The KYC process is part of the growing regulation of the financial industry that began with the Money Laundering Control Act of 1986 (see USA (1986)) and has been growing extensively since in the form of further, ongoing regulation aimed at precluding either money laundering or the funding of terrorist activity (see USA (1988, 1992, 1994, 1998, 2001, 2004)). Financial institutions are obliged by regulators to onboard their customers before conducting any activity with them, in order to avoid working with customers that pursue either of the aforementioned illicit activities. The KYC process consists of an exchange of documents between the customer and the financial institution that intend to work together. The process includes the collection of basic identity information from all beneficiaries to check for illicit activity and "politically exposed persons."[2] The process also includes risk management with regard to onboarding new customers, the monitoring of transactions, and specific customer policies for banks. The process is costly for financial institutions and may expose them to large fines if it is not conducted in accordance with the existing regulations (e.g., HSBC was fined USD 1.92 billion when it was discovered that Mexico's Sinaloa cartel and Colombia's Norte del Valle cartel had laundered USD 881 million through the bank (Viswanatha and Wolf, 2012), and ING Bank paid USD 619 million in fines for violating sanctions against a variety of countries (Freifeld, 2012)).

The KYC process is initiated when a customer intends to work with a financial institution. Chronologically, the customer and the financial institution agree on the terms of a relationship. Then, the customer sends the required documents to the financial institution in order to enable the institution to conduct the KYC verification process. The financial institution analyzes the documents and generates an additional, internal document that serves as the certification that assures regulators that this customer has been either validated or rejected and that the KYC process has been properly conducted. This process is repeated every time the customer intends to work with a new financial institution. In the current setting, every time a customer initiates a relationship with a financial institution the costs of the KYC verification process recur. Figure 1 shows an example case that illustrates the process that occurs when a customer intends to work with three different financial institutions. This example case shows how, for this single customer, the exchange of documents and the core KYC validation must be undertaken three times, such that the total costs that are generated by this customer are three times those of a single KYC process. At this point, it is important to differentiate between the "core KYC verification process", which is the minimum KYC verification that all financial institutions are obliged by law to conduct, and additional, bank-specific processes. While further documentation can be asked for by each financial institution to create an "additional aura of information" for every customer, our solution focuses solely on the core KYC verification process, which is that shared by all the financial institutions in a jurisdiction.

The growth of regulation and changes to technology, as well as the financial crisis of 2007, have created opportunities for companies, working in

[2] Broadly speaking, a "politically exposed person" is one who has been entrusted with either a prominent public function or a prominent function in a state-owned enterprise or international organization, either at home or abroad. Such individuals must be identified during the process given their particular ability to exert considerable influence.

a field referred to as “regtech”, that aim to use technology to improve the implementation of regulations. The term “regtech” comes from the combination of the words “regulation” and “technology”. These opportunities are especially significant within the domain of KYC (see Memminger et al (2016) and Arner et al (2016)). Arasa and Ottichilo (2015) conduct an analysis of the cost of KYC based on the complexity level of the compliance required for the case of commercial banks in Kenya, establishing four variables that explain 78.3 percent of the compliance requirements. Soni and Duggal (2014) look into using big data analytics to reduce risk for institutions conducting the KYC process. Colladon and Remondi (2017) work on different approaches to using cluster analysis over a network of customers and potential customers to identify suspicious financial operations and potentially criminal activities. They do so by mapping relational data and using predictive models over an internal transactions database involving data from over 33,000 financial operations. A survey of the latest regulatory requirements and a history of KYC and AML processes can be found in Ruce (2011). KYC can be improved by, for example, improving auditors’ effectiveness in assessing KYC and AML practices. A case study in the context of Luxemburg is provided by Smet and Mention (2011) and reveals that audit effectiveness could be increased and information asymmetries reduced by an ISO standard for an internal control assessment model for KYC. The current paper aims to deliver an additional improvement by using DLT to reduce the aggregate cost of the KYC process and distribute these lower costs proportionally among the financial institutions participating in the system. Tackling the cost of the KYC process from the aggregated perspective (i.e., as the sum of the individual costs of each financial institution) and using DLT to reduce this aggregate cost is the main contribution of this paper.

3 Blockchain Technology

DLT, such as blockchain technology, has gained prominence thanks to the widespread use of the cryptocurrency Bitcoin. Bitcoin, introduced by Nakamoto (2008), was the first working cryptocurrency that was not owned by a central authority. While DLT was originally used to provide a new way of creating money and transferring it via the Internet, the technology can also be used to run and govern decentralized systems by means of smart contracts. Smart contracts are computer protocols that facilitate, verify, or enforce predefined clauses whenever a set of conditions is given. As described by Szabo (1997), the intention of using smart contracts is to embed them in a whole

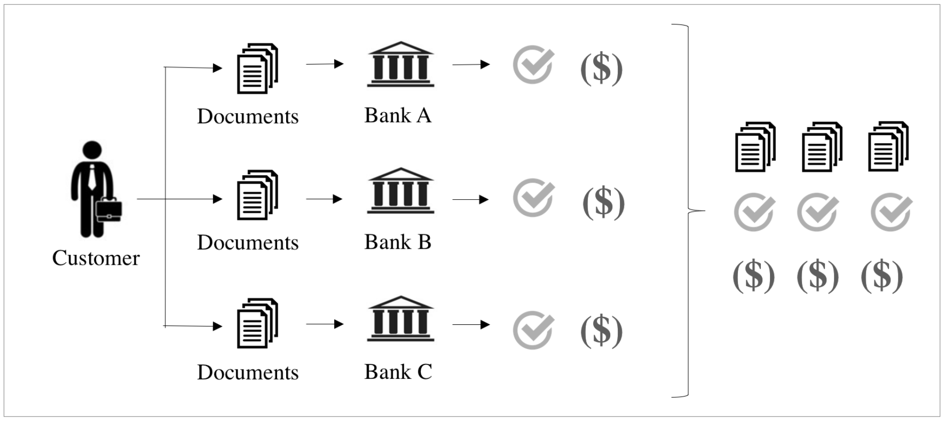


Figure 1: Current process and cost structure of KYC

range of properties that are valuable and controlled by digital means. Since Nakamoto's seminal work (Nakamoto (2008)), new instances that propose the use of DLT for a range of novel purposes have emerged. One of these is "Ethereum", which is a platform upon which whole decentralized applications may be run (see Wood (2016)). Many papers, including Peters and Panayi (2015) and Harvey (2016) discuss the blockchain from a technical perspective.

While transactions in the Bitcoin blockchain can include small scripts that define output spending conditions, such as the requirement that a transaction be signed by two keys instead of one, the Ethereum blockchain can be seen as a Turing complete virtual machine that can run code in several programming languages and therefore run the smart contracts stored in it (see Glaser (2017)).

Glaser (2017) provides a solid ontological development of blockchain systems concepts and defines a common set of blockchain components and relationships. This analysis serves as a framework and basis for assessing the implications of blockchain solutions in an academic or economic context. Further, it introduces the perspective of a pervasive decentralization of multiple layers of digital infrastructure by blockchain technology. Specifically, Glaser (2017) defines and describes two layers of code —namely, the fabric layer and the application layer. The term fabric layer denotes the system's code base, which embraces communication, the public-key infrastructure, the software that constructs and maintains the database, and the execution environment of the system. Whoever develops and maintains the fabric layer controls the functioning of the system. Ultimately, the fabric layer defines the governance type of the system, which can be the only dimension of the fabric layer, and that can be public, permissioned, or hybrid. Nevertheless, and as described by Glaser (2017), one important characteristic of blockchain systems is that they do not allow for a differentiation between users and user management modules, which implies that all the users have complete transparency when reading the transactions and the smart contract code deployed.

The application layer comprises the application logic of the services implemented in the form of smart contracts. The application layer encompasses three dimensions —namely, the ecosystem closedness, the value linking, and the market type. The closedness of the ecosystem refers to the extent to which the system needs to interact with other structures that are outside of the blockchain-based framework —that is, with other trusted interfaces. Since the decentralization of control ends at the boundaries of the blockchain-based system, the more closed the system is, the higher the leverage of a blockchain-based solution. The value linking of the system refers to the intrinsic value of the tokens that are exchanged between parties within the system. Glaser (2017) suggests four possible ways in which value is assigned to the tokens of a system —namely, being the token a community currency, being seen as debt or equity by the participants of the system, being backed by a commercial bank, or being backed by a central bank. The last dimension of the application layer is market type, which describes the nature of the market in which the blockchain-based solution is framed.

The European Security and Markets Authority (2016) sets out the possible benefits of DLT applied to securities markets, discusses the possible shortcomings of and challenges to those benefits, and analyzes the relevant regulatory framework, with a focus on the main EU legislation relevant to potential applications of DLT in securities markets. While the Authority focuses on the securities market, it provides a DLT-solutions governance framework that can be very similar to the governance framework required by the solution proposed in this paper. Specifically, it suggests that for the interbank context of securities markets, a permission-based system can be of great value. Further, the Authority claims that such a system would allow for governance of the interaction between the system's participants, paying special attention to the liabilities of each participant, correction mechanisms, and even penalties in the case of infringement of the rules.

The European Central Bank (2012) defines and classifies virtual currency schemes based on their observed characteristics. Depending on the interaction of the virtual currency schemes with traditional money and the real economy, the Bank classifies them into three types: Type 1, which refers to closed virtual currency schemes, which operate in the same way as do virtual currencies used in online gaming; Type 2, virtual currency schemes with a unidirectional flow (usually an inflow), meaning that there exists a conversion rate for purchasing the virtual currency; and Type 3, virtual currency schemes that have bidirectional flows.

The World Economic Forum (2016) analyzes the current phase of the disruptive innovation work that is being conducted in terms of DLT in the financial sector, first looking at how blockchain can reshape financial services, and then studying the role of financial institutions in building digital identity. The Forum (2016) concludes that DLT can enable the design of new systems or improve existing ones, by automating processes, reducing settlement time, reducing costs, reducing operational risk, providing central authority disintermediation, and offering real-time settlement.

(Egelund-Müller et al, 2017) look into the construction of an automated financialsystem, with multiple counterparties, that can run a variety of complex financial derivatives, including settlement, directly on DLT.

4 Design Science for KYC Optimization

According to (Hevner et al, 2004, p. 77), the objective of design science research (DSR) is to produce a technology based solution —in the form of a viable artefact —that solves a relevant business problem. In the context of a hackathon organized at the IT University of Copenhagen, we collaborated with the Nordic financial services group Nordea Bank AB to study the inefficiencies and costs related to the KYC process, and analyzed if this process could be improved by means of a DLT-based solution. During these four days we were confronted with the aforementioned reality of KYC inefficiencies, and transformed the existing problematic into the following research question:

"Can a DLT-based solution reduce the cost of the KYC process for financial institutions and improve the customer's experience?"

In order to answer the research question and to design an effective artefact that solves the problem at hand within the corporate and regulatory context, we followed Hevner et al (2004)'s DSR approach and focused on its three components (environment, IS research, and knowledge base). To strengthen the utility, quality, and efficacy of the proposed solution, we also considered the DSR process based on Peffers et al (2007)'s approach, which synthesizes design science processes from Information Systems (IS) and other disciplines. This process is subdivided into five sub-steps: problem identification, objective definition, design and refinement of the artefact, demonstration of the artefact, and evaluation of the artefact. The last three steps of the process need to be repeated recursively in a loop in order to gather feedback from the environment and to refine the artefact according to that feedback. Both the approach and the process are summarized in Figure 2.

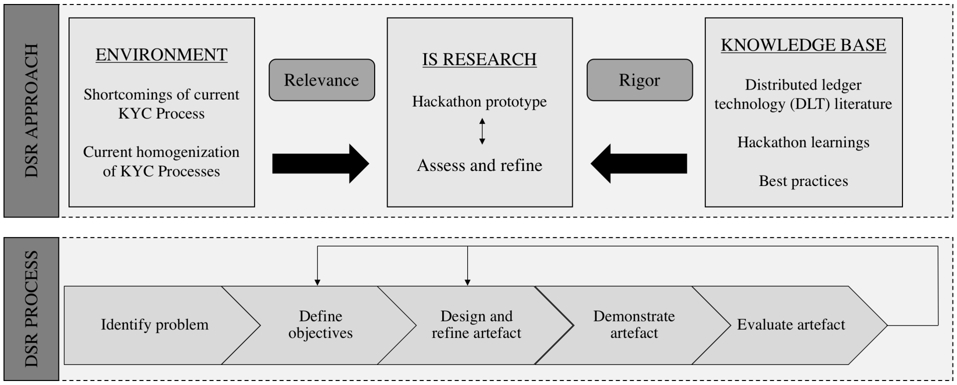


Figure 2: DSR approach and DSR process. Source: authors' own illustration adapted from Hevner et al (2004) and Peffers et al (2007).

Nordea Bank AB, representing the corporate environment, expressed the need for improvement in the KYC process. They provided us with information concerning the applied difficulties of conducting the process and pointed out its main pain sources. This enabled us to identify the problem and define our objective (previously formulated in the form of our research question): use a DLT-based solution to reduce the cost of the KYC process for financial institutions and improve the customer's experience. In order to better understand the environment, we researched the existing KYC literature, paying special attention to efforts made in recent years to homogenize the KYC process and increase its efficiency without compromising security. Further, we held various exchanges with experts in the field (lawyers, practitioners, and experts) regarding best practices in KYC. During these exchanges, it became clear that the system proposed would need to fulfil three conditions if it was to be accepted by the participants. First, it would need to enable its users to obtain a tamper-proof record of the KYC process in the case of conflict. Second, it would have to reduce the costs of the current KYC process and distribute the remaining costs in a proportionate manner among the participants of the system. Third, the system would need to not compromise the responsibility of banks with regard to conducting the KYC process. The combination of the environment's needs and our knowledge base constituted the grounds for our IS research, which yielded the first version of our artefact, a version that we continued to refine over several months based on ongoing dialog with, and feedback on the artefact from, KYC practitioners. With the problem identified and our objective defined (see above), the first design and refinement phase of the artefact was conducted, taking into account the feedback and validation of KYC practitioners, as well as the insights with regard to DLT from our knowledge base and the KYC experience from the environment. The first demonstration of the artefact took place during the Nordic Blockchain Summit, at which it was awarded first prize, receiving the majority of the votes of an audience of over 300 practitioners from the senior corporate management level. The first evaluation phase involved various informal working sessions with KYC practitioners who studied the artefact in terms of its relevance and viability, which helped us to learn more about the specific requirements of the participants.

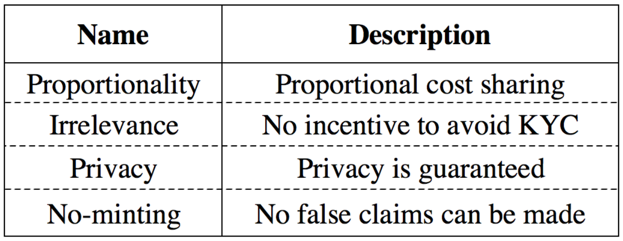
After the first design, refinement, demonstration, and evaluation phases, we undertook a second loop of refinement, demonstration, and evaluation, following the DSR process described in Figure 2. The second loop incorporated the feedback of five senior executives from the banking sector, a lawyer, and two senior government officials, with whom we conducted several working sessions to explore various implementation possibilities of the solution here proposed. Their feedback was related to the need for interbank collaboration and for cooperation with the national regulator, as well as the need to launch the process in a single, relatively small country (that can amend the required regulations efficiently and quickly), to ensure that the system functions correctly. This feedback round made us aware of the need to initially propose the solution at a national level, moving on to a solution that would encompass a range of countries only later. From these working sessions, we also learned about the central role of the national regulator as the cornerstone of such a DLT-based solution, about the need to identify the individuals involved at each step of the KYC approval process, and about the importance of keeping all the documents of a specific customer on a secure local storage facility with only the hashes of each document stored on the DLT (in order to facilitate the tracing of past activity while ensuring that banks still know their customers and can effectively protect customer privacy with regard to cyber attacks). These points were influential in our decision to assign to the national regulator the role of maintaining the system.

5 The Redefined KYC Process

The IS suggested in this paper to solve the current inefficiencies of the KYC process relies on the following three assumptions: First, a group of financial institutions, working in the same country and therefore obliged to respect the same KYC regulations, agrees on the standards for granting core KYC verification to a customer. Second, all the financial institutions that collaborate in the system agree on the average costs of conducting a core KYC verification process. This cost might of course depend on the complexity of each individual customer, based on predetermined parameters (e.g., client size, volume of documents exchanged, etc.). Third, the national regulator maintains the system and approves financial institutions to work with the system in order to conduct a more efficient and transparent KYC verification process. These three assumptions are necessary to ensure a correct incentive structure across the participating financial institutions.

Further, we define a set of four conditions that must be fulfilled by the artefact. It must ensure the proportional sharing of the cost of conducting the core KYC verification process; maintain the privacy standards of the KYC process as they are today; ensure that no institution can claim compensation without conducting that core process; and ensure that no institution can become a free rider and avoid paying for using the information generated by other member institutions. The proportionality condition ensures that the costs are shared proportionally. The irrelevance condition ensures that the financial institution that conducts the core KYC verification process does not have an incentive to prefer that another institution conducts the core KYC verification process and vice versa. The privacy condition ensures that the financial institutions that work in the system cannot know with which other financial institutions the customer is working, unless the customer reveals that information (privacy is required among financial institutions). The no-minting condition ensures that no financial institution can simulate having conducted a core KYC verification process in order to be compensated by other institutions for work that it has not done. These conditions are summarized in Table 1.

The suggested artefact is composed of two parts. The first part is a permissioned database that stores the documents that require a certain privacy.



Name	Description
Proportionality	Proportional cost sharing
Irrelevance	No incentive to avoid KYC
Privacy	Privacy is guaranteed
No-minting	No false claims can be made

Table 1: Conditions for ensuring the viability of the system

The second part is a distributed ledger that serves as an immutable record and clearing system via which to proportionally distribute the costs of the KYC process among the participating institutions. The system is held and managed by the regulator, who enables the database and the DLT infrastructure. This implies that the national regulator develops and maintains the fabric layer and therefore plays a central role in the system. The clearing itself, however, is conducted via the smart contract, which comes along with very low clearing costs for this solution. The artefact works as follows.

1. A number $k \geq 3$ of financial institutions and the national regulator agree to interact with the artefact and set the average price m of conducting a core KYC verification process. The regulator establishes a digital currency with a fixed exchange rate against the national currency. This automatically assigns value to the token used in the system. In terms of the abovementioned European Central Bank (2012) classification, this system would be framed as a Type 3 virtual currency scheme. Each financial institution can purchase digital currency in exchange for national currency, such that it can later on compensate other member financial institutions for the verifications that they conduct. The purchased digital currency can be distributed across as many different accounts as each financial institution desires. Since the system is run by the regulator, no financial institution can know to which financial institutions the other accounts belong. Only the regulator is aware, with certainty, of the activities of each financial institution.

2. Whenever customers approach a member financial institution to be validated in terms of KYC for the first time, they are granted a new account (with a public and a private key) through the systems interface. For the sake of brevity, we refer to the first financial institution that conducts the core KYC verification for a customer as the “home bank”. Once customers have been granted an account in the system, they can share with the home bank their public key and the documents that must be analyzed. The exchange

of these documents occurs outside of the distributed ledger to protect the privacy of the customer. The home bank will keep these documents in its local database. Once the bank decides to validate or reject a customer, it stores a digitally signed document in the smart contract of this customer and this includes the result of the core KYC verification process (*verified* or *rejected*). Additionally, the hash of each of the documents submitted by the customer, documents that have been used for the verification, is also stored by the home bank on the distributed ledger. Once the validation has been conducted, the home bank creates a "document package" for the customer, which contains the documents submitted by the customer and that have previously been hashed, as well as the digitally signed document that summarizes the KYC verification process and includes the result of the core KYC verification. This document package is stored in the bank's local database as well as in the permissioned database managed by the regulator. At this stage, only the customer and the home bank have the documents package. Further, the home bank creates a smart contract for this customer, a contract that contains a list of the public keys of the wallets of the financial institutions that have checked that the status of this customer in terms of KYC has been verified and that have paid their corresponding fraction of the verification costs. We call this list the "list of onboarding institutions". At the time of its creation, when a customer only works with the home bank, the list of onboarding institutions only contains the public key of the account that the home bank has used to interact with this customer. This list can later be enlarged as the customer interacts with further institutions. We suggest that each bank uses a single, unique, one-payment-only account to interact with each customer, since this will later on protect the privacy of financial institutions and customers.

3. Whenever customers approach an institution other than the home bank with the intention of working with it, they can share with it their public key and key and the address of the original smart contract in which the home bank wrote the result of the core KYC verification process. Further, they can grant this institution access in the permissioned database to the documents package previously created by the home bank, such that it too can read them and validate the customer. Further, by reading the smart contract, the new financial institution can see how many institutions have worked with the customer so far, since it can see how many public keys appear in the list of onboarding institutions. To be added to this list, a financial institution has to pay the proportional part of the average price m of conducting a core KYC verification process. Specifically, this institution has to pay $\frac{m}{k}$ to the smart contract. Note that $k-1$ is the number of institutions that have worked

with this customer so far (i.e., $k-1$ is the number of institutions that are listed in the list of onboarding institutions). The smart contract then sends the compensation that it has received, divided into equal parts between the $k-1$ institutions that had previously worked with this customer, and adds the public key of the account from which it has received the payment to the list of onboarding institutions. The payment is made in the cryptocurrency issued by the regulator.

4. This mechanism ensures that all the financial institutions that work with one given customer share the costs of the core KYC verification process proportionally; that is to say, if the k-th institution that starts working with a customer always contributes with $\frac{m}{k}$ and this contribution is distributed in equal parts among the accounts of the other $k-1$ institutions, all the institutions that work with the customer end up paying the same fraction of the average price m of conducting a core KYC verification process. It is easy to see that for $k=1$ only the home bank works with the customer and that it bears the full average cost m of conducting a core KYC verification process, since no other institution is compensating it for the work conducted, which is worth m. For the case in which $k=2$, the second financial institution to join pays $\frac{m}{2}$ to the smart contract, which automatically sends this compensation to the home bank, such that both institutions bear a cost equal to $\frac{m}{2}$. Let us assume now that this system works for a number $k \geq 2$, such that the k-th institution pays $\frac{m}{k}$. So far, each of the other $k-1$ institutions has paid $\frac{m}{(k-1)}$ and now receives an amount equal to $\frac{m}{k(k-1)}$ from the last institution to join. Hence, the cost for each institution equals $\frac{m}{k-1} - \frac{m}{k(k-1)} = \frac{m}{k}$.

The smart contract contains the documents' hash codes, the public key of the home bank, the certificate of approval, which conveys that the customer has been validated, and an array called "onboarded" with all the public keys of the financial institutions that have paid the proportional compensation amount to the home bank.

This system ensures that the core KYC process only has to be undertaken once, by the first institution with which a customer intends to work, but that its result can be used by as many financial institutions as required by the customer. This specific setting shows how, for a customer that works with k financial institutions, the exchange of documents and core KYC verification need only be undertaken once (and not k times as is the case in the current setting). Furthermore, the total cost of conducting the core KYC verification for one customer is now the cost m of one single KYC (and not k*m, as in the current practice).

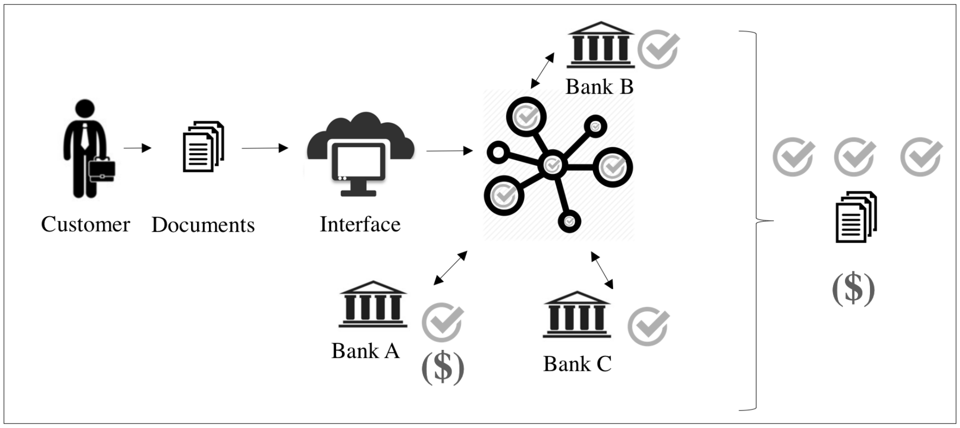


Figure 3: Proposed work flow and cost structure of KYC after the implementation of the artefact

Figure 3 illustrates the same example case as that presented in Figure 1, but this time following the introduction of the proposed system. The system enables the same customer to work with the same three financial institutions, but now the exchange of documents and the core KYC verification process only occur once and the costs are reduced to a third.

This system fulfills the four previously defined conditions: proportionality, irrelevance, privacy, and no minting. With regard to privacy, since each financial institution only uses one account for each customer, and it is therefore not possible to identify which institution is behind which public key, privacy, for customers and financial institutions, is ensured. Only if one customer would work with all the institutions in the system would all the institutions be able to infer that this was the case. However, since financial institutions use only one account per customer, their privacy would still be guaranteed with regard to the rest of the customers. The no-minting condition is fulfilled, since only by paying can an institution be added to the onboarding institutions list of a customer that approaches it. Since the action of compensating other institutions for the core KYC verification process that has been conducted can only be triggered by a real customer approaching an institution, no institution has an incentive to fake smart contracts claiming that it has conducted a core KYC verification process, since in such a case there would exist no genuine customer behind such a process that would subsequently approach another institution and ask to be verified.

6 Implementing the Redefined KYC Solution

In this section we discuss the implementation considerations of the DLT-based KYC solution previously described. It is important to note that the implementation of such a system would have significant implications for the financial sector and that it would therefore need to be carried out in close coordination with the regulator. Further, many of the dimensions of the system would depend on specific national guidelines and legislation. Hence, in this section we discuss both the suggested system and two variations on it that offer different degrees of centralization and thus make possible its implementation. We also discuss alternative designs and look into the challenges and benefits of those designs.

6.1 Design of a KYC Solution

The system proposed in Figure 4 explains the new KYC process using the example of a customer that approaches two financial institutions. In a first step, the customer approaches the home bank and provides the required KYC documents for verification. The home bank uses the system's application (which is installed at each of the participating documents onbanks) to handle the process of document exchange with the customer outside of the distributed ledger and to store these documents in its local database. When any document is processed by the home bank, the hash of the document is stored on the distributed ledger. Once the home bank has validated the customer, it can create the abovementioned document package, which contains all the documents that have been used (and previously hashed) to grant the verification status, as well as the digitally signed document that grants verification to this customer. Later on, the customer can provide access to this document package to any other institution with which it intends to work. Hence, the next institution that needs to validate this customer in terms of KYC can use the local client application and communicate with the smart contract of the customer in order to obtain the customer's status, inscribe itself in the list of onboarding institutions, and handle the necessary payment over the blockchain as described in the previous section. Further, since this institution has been granted access to the document package by the customer, it can store a copy of it locally on its own database.

In the proposed solution, the regulator is assigned a central role as a trusted third party (TTP) and owner of the "fabric layer". This could represent a possible shortcoming of the system if —for example —the regulator were corrupt, or compromised by hacking or by insider fraud. This is indeed

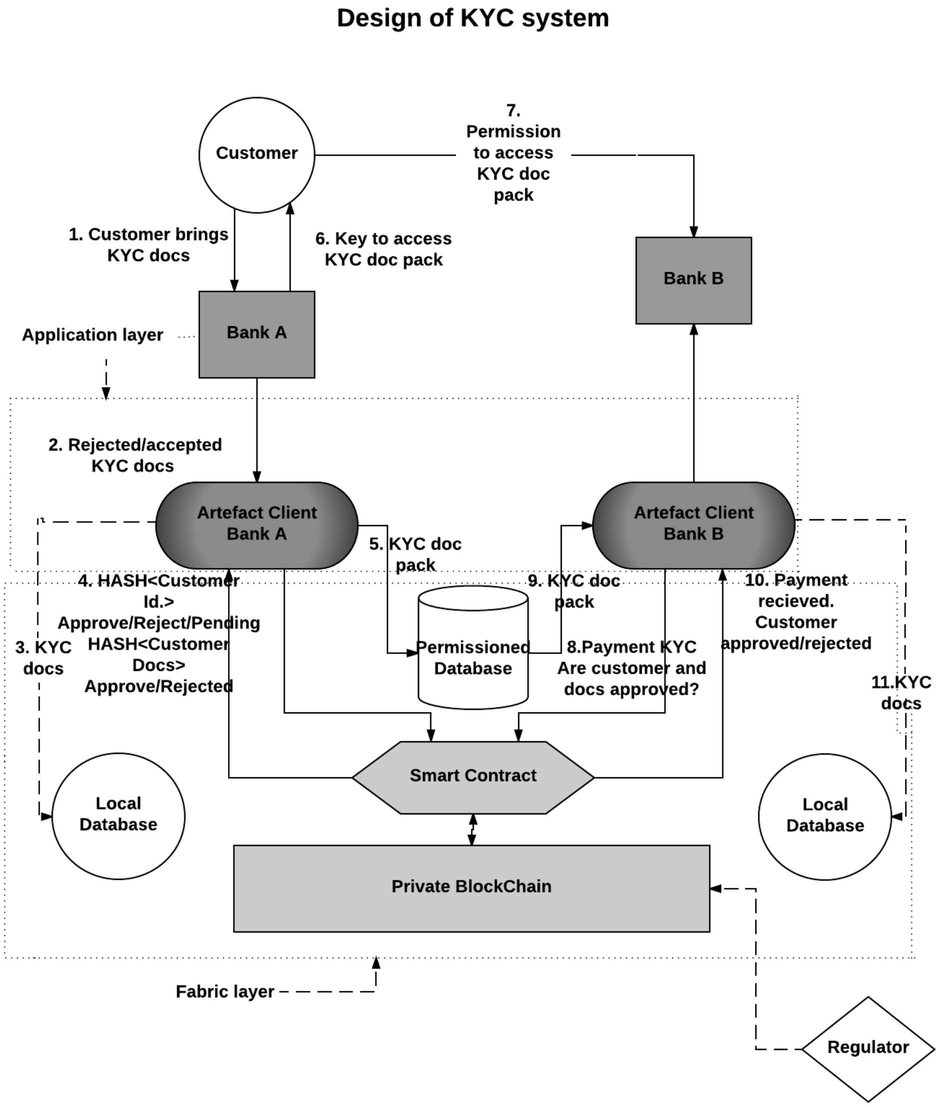

Design of KYC system
Customer
7. Permission to access KYC doc pack
1. Customer brings KYC docs
6. Key to access KYC doc pack
Bank B
Application layer
Bank A
2. Rejected/accepted KYC docs
Artefact Client Bank A
5. KYC doc pack
Artefact Client Bank B
4. HASH<Customer Id.> Approve/Reject/Pending HASH<Customer Docs> Approve/Rejected
9. KYC doc pack
10. Payment recieved. Customer approved/rejected
3. KYC docs
Permissioned Database
8.Payment KYC Are customer and docs approved?
11.KYC docs
Smart Contract
Local Database
Local Database
Private BlockChain
Fabric layer
Regulator

Figure 4: Design of the KYC solution

an aspect that can be further analyzed in the future. In order to mitigate this potential shortcoming to a certain extent, the TTP characteristics described by Lee et al (2016) could be incorporated.

6.2 Decentralized KYC Solution

The solution proposed in the previous subsection can be further decentralized with the following modifications. First, if the DLT part of the solution were implemented directly on the Ethereum network rather than using a private blockchain, any attempt to change the information on the blockchain would be made more difficult due to the existence of a large mining community that is harder to corrupt. Second, the regulator could be removed from the system, thus precluding the risk of there being a party that has an unlimited view of the system. Last, some further efficiency could be introduced by storing the data only at the financial institution that has actually approved the customer. This solution is shown in Figure 5. While we acknowledge these benefits, our discussions with experts indicate that in most Western countries the risk of a corrupt regulator is considered low when weighed against the benefit of the higher financial stability that would result from the regulator's ability to easily and routinely check the KYC process. Furthermore, storing the documents locally ensures that any bank that works with a client would check of the KYC documents whenever it wished. In our proposed design we have used a private distributed ledger and not a public one. This decision was based on the feedback received from the finance executives consulted during the DSR process, who stated that banks would not be comfortable having customers' private information available on a public distributed ledger (even if only hash code values of documents and the key to decrypt the customer document package were to be kept on a public ledger). This is understandable, as potential bugs in the smart contract or reverse engineering of the smart contract bytecode could lead to the risk of exposing information unintentionally. Luu et al (2016) scan 19,366 smart contracts on Ethereum and find vulnerabilities in 8,833 of them. The stated concerns of the finance executives consulted are, then, well grounded. Further, the whole compensation scheme that enables the cost reduction and cost sharing within the system is only possible thanks to the use of DLT.

A more mature DLT would allow for a ledger in which stored documents can be held completely privately. This would make possible a decentralized, permissioned database held on a blockchain. In such a system, the document package would only be stored on a distributed ledger, and not on a central database managed by the regulator. The projects R3 Corda and Hyperledger are moving in this direction. While these projects are not mature enough currently, they may well be in the near future.

6.3 Centralized KYC Solution

It is possible to benefit from cost sharing during the KYC process by using a different, fully centralized KYC arteafct. This would require only one party being allowed to approve or reject customers. One such centralized solution would be to transfer the entire KYC responsibility to one specialized entity or a regulator-operated KYC office. In such a design, the customer would need to be authorized by the entity and, subsequently, each bank that wanted to work with that customer would obtain a permission to do so from the centralized authority. This solution is shown in Figure 6, and while itis unlikely to be adopted as it creates an additional cost for the regulator and in essence frees banks from the responsibility of knowing their customers, there are some significant benefits to be gained from such a solution. The main benefit is that by removing the costs of KYC from banks (and other financial institutions) we reduce significantly the cost of forming a new financial entity and, in this way, open the market up to increased competition. Furthermore, this reduction in costs for banks would lead to lower fees for customers and lower costs for doing business in a given country. That in turn would benefit a country that uses a centralized KYC solution as that country would be perceived as being open for business and competitive without necessarily compromising AML or KYC requirements.

6.4 The Use of Distributed Ledger Technology

Having presented a solution, it is worth considering why the use of DLT represents an improvement compared to other possible technologies. First, there would be improvements in terms of auditing and tracking. This is advantageous for the national regulator since it provides a clear record of the information that financial institutions verify prior to the opening of accounts, and could serve as a single point of truth should disagreement occur. And the immutable nature of the record created by DLT-based solutions cannot be matched by other technologies. Second, the proposed system allows collaboration between financial institutions that do not necessary trust one another. Specifically, given that financial institutions compete for customers' assets and accounts, only a system that allows for anonymous collaboration —such as anonymous compensation and anonymous document sharing —would gain the support of financial institutions. Third, one of the major contributions of the solution proposed here is that an institution can be anonymously and proportionately compensated by others for the efforts conducted to verify a customer. This is only possible due to the features of the distributed ledger, which allow institutions to communicate with one another without revealing their identities but ensure that each institution abides by all relevant regulations at all times. Fourth, it is important to note that the system proposed here —irrespective of the technology used to enable it —is, in essence, a system for interbank collaboration. Since financial institutions are studying broader interbank collaborations based on DLT —such as the R3 project —it seems logical to propose a system such as the one presented here, which already takes core DLT features into account, such that it can, in the future, be integrated into a broader DLT-based framework. Last, and taking into account that such a novel system would in any case need a clearing instance to settle the compensations, DLT eliminates high central authority fees. All in all, the solution proposed here from DLT for the following reasons: the application of this technology allows for the automation of a process, increases the information available if a dispute should occur, reduces settlement time compared to other technologies, and reduces business costs.

7 Conclusion

This paper has suggested an IS to reduce the aggregated cost of KYC in a jurisdiction by means of DLT. The main efficiency gain that this IS proposes is the avoidance of the same tasks being duplicated by different financial institutions. Additionally, this paper has shown how it is possible to distribute the costs of the core KYC verification process proportionally among those financial institutions, solutions that require the verification process be carried out for one given customer, and has defined a series of conditions that the IS in question needs to fulfil in order to ensure the correct incentive structure for the participating institutions. The maximum total cost saving per customer generated by the proposed IS can be measured as $\sum_i m_i * (k_i - 1)$, where m_i is the cost of conducting a full core KYC verification for a customer i, and k_i is the number of financial institutions that conduct business with customer i. This implies that the monetary savings brought about by the proposed IS and the increased efficiency that it would deliver for both customers and institutions are significantly affected by the number of financial institutions that participate in the system. The proposed IS has emerged from the application of design science research to the problems of high costs for financial institutions and the low satisfaction of customers when conducting a core KYC verification process. The fact that the smart contracts in which the information is stored would be owned by the customers and not by the participating institutions already addresses the paradigm shift taking place with regard to consumer data in light of the General Data Protection Regulation (GDPR), which will come into force in 2018 (European Commission (2016)). For example, a simple extension of the system could oblige the client application running at each bank to regularly check in order to detect if a customer has decided to no longer work with the bank and ensure that that customer's private documents are deleted. Performing a core KYC verification process on a distributed ledger has many intersections with ongoing research in the area of digital identity in distributed ledgers. One question that arises here is that of the location in which customers' sensitive documents would be stored. In the proposed IS, all the information is stored locally by each bank, as well as in a permissioned database maintained by the regulator. This is primarily due to the high cost of storage on the Ethereum platform on which the artefact was first designed. It is possible to conduct other designs based on permissioned, contractually based solutions such as R3CEV's Corda or Monetas, both of which are currently generating a lot of interest. Corda and the Ethereum blockchain have similarities, but the former is — in its essence — the combination of a distributed database and a Java Virtual Machine, enabling parties on the network to execute bilateral transactions involving sensitive information that is not revealed to the public. These kinds of solutions could offer new approaches to providing distributed but private document exchange between customers and financial institutions that include storage possibilities for larger documents. However, solutions such as Corda are still in their early stages of development and privacy with regard to the customer data that is shared in such a system is a concern that needs to be thoroughly addressed.

Regardless of the chosen approach to using DLT, be it a distributed database or a private, restricted, or public blockchain, our research suggests many opportunities to increase efficiency in the financial system. More specifically, a significant reduction in costs for the participating institutions and an improved experience for customers could both be delivered by such a system. Furthermore, the system would — thanks to the decreased regulatory costs of KYC — lower the barriers to operating a financial institution, thus opening the financial market up to further competition.

References

Arasa R, Ottichilo L (2015) Determinants of know your customer (kyc) compliance among commercial banks in kenya. Journal of Economics and Behavioral Studies 7(2):162–175

Arner D, Barberis J, Buckly R (2016) The emergence of regtech 2.0: From know your customer to know your data. Journal of Financial Transformation 44:79–86

Colladon AF, Remondi E (2017) Using social network analysis to prevent money laundering. Expert Systems with Applications 67:49–58

Egelund-Müller B, Elsman M, Henglein F, Ross O (2017) Automated Execution of Financial Contracts on Blockchains. URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2898670

European Central Bank (2012) Virtual Currency Schemes. URL https://www.ecb.europa.eu/pub/pdf/other/virtualcurrencyschemes201210en.pdf

European Commission (2016) Reform of EU data protection rules - European Commission. URL http://ec.europa.eu/justice/data-protection/reform/index{_}en.htm

European Security and Markets Authority (2016) The Distributed Ledger Technology Applied to Securities Markets. URL https://www.esma.europa.eu/sites/default/files/library/2016-773_dp_dlt.pdf

Freifeld K (2012) ING to pay $619 million over Cuba, Iran sanctions. URL http://www.reuters.com/article/us-ing-sanctions-idUSBRE85B12I20120612

Glaser F (2017) Pervasive decentralisation of digital infrastructures: A framework for blockchain enabled system and use case analysis. In: Proceedings of the 50th Hawaii International Conference on System Sciences (HICSS 2017) (forthcoming)

Harvey CR (2016) Cryptofinance. URL https://ssrn.com/abstract=2438299

Hevner AR, March ST, Park J, Ram S (2004) Design science in information systems research. MIS Q 28(1):75–105, URL http://dl.acm.org/citation.cfm?id=2017212.2017217

Lee K, James J, Ejeta T, Kim H (2016) Electronic voting service using block-chain. The Journal of Digital Forensics Security and Law

Luu L, Chu DH, Olickel H, Saxena P, Hobor A (2016) Making smart contracts smarter. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, ACM, pp 254–269

Memminger M, Baxter M, Lin E (2016) Banking Regtechs to the Rescue? URL `www.bain.com/Images/BAIN_BRIEF_Banking_Regtechs_to_the_Rescue.pdf`

Nakamoto S (2008) Bitcoin : A Peer-to-Peer Electronic Cash System. URL `https://bitcoin.org/bitcoin.pdf`

Peffers K, Tuunanen T, Rothenberger MA, Chatterjee S (2007) A design science research methodology for information systems research. Journal of management information systems 24(3):45–77

Peters GW, Panayi E (2015) Trends in crypto-currencies and blockchain technologies: A monetary theory and regulation perspective. URL `https://arxiv.org/abs/1508.04364`

Ruce P (2011) Anti-money laundering: The challenges of know your customer legislation for private bankers and the hidden benefits for relationship management ('the bright side of knowing your customer'). The Banking Law Journal 128(6):548–564

Smet DD, Mention A (2011) Improving auditor effectiveness in assessing kyc/aml practices. Managerial Auditing Journal 26(2):182–203, DOI 10.1108/02686901111095038

Soni A, Duggal R (2014) Reducing risk in kyc (know your customer) for large indian banks using big data analytics. International Journal of Computer Applications 97(9)

Szabo N (1997) Smart contracts: Formalizing and securing relationships on public networks. Expert Systems with Applications 2(9)

Thompson Reuters (2016) Know your customer (kyc) independent survey

USA (1986) The money laundering control act of 1986 (public law 99-570)

USA (1988) The anti-drug abuse act of 1988, pub law 100–690, 102 stat. 4181

USA (1992) Annunzio-wylie anti-money laundering act of 1992

USA (1994) The money laundering suppression act of 1994

USA (1998) The money laundering and financial crimes act (1998)

USA (2001) Usa patriot act of 2001 (public law 107-156)

USA (2004) Intelligence reform & terrorism prevention act of 2004

Viswanatha A, Wolf B (2012) HSBC to pay $1.9 billion U.S. fine in money-laundering case. URL http://www.reuters.com/article/us-hsbc-probe-idUSBRE8BA05M20121211

Wood G (2016) Ethereum: A secure decentralised generalised transaction ledger. URL http://gavwood.com/paper.pdf

World Economic Forum (2016) The Future of Financial Infrastructure. URL http://www3.weforum.org/docs/WEF_The_future_of_financial_infrastructure.pdf

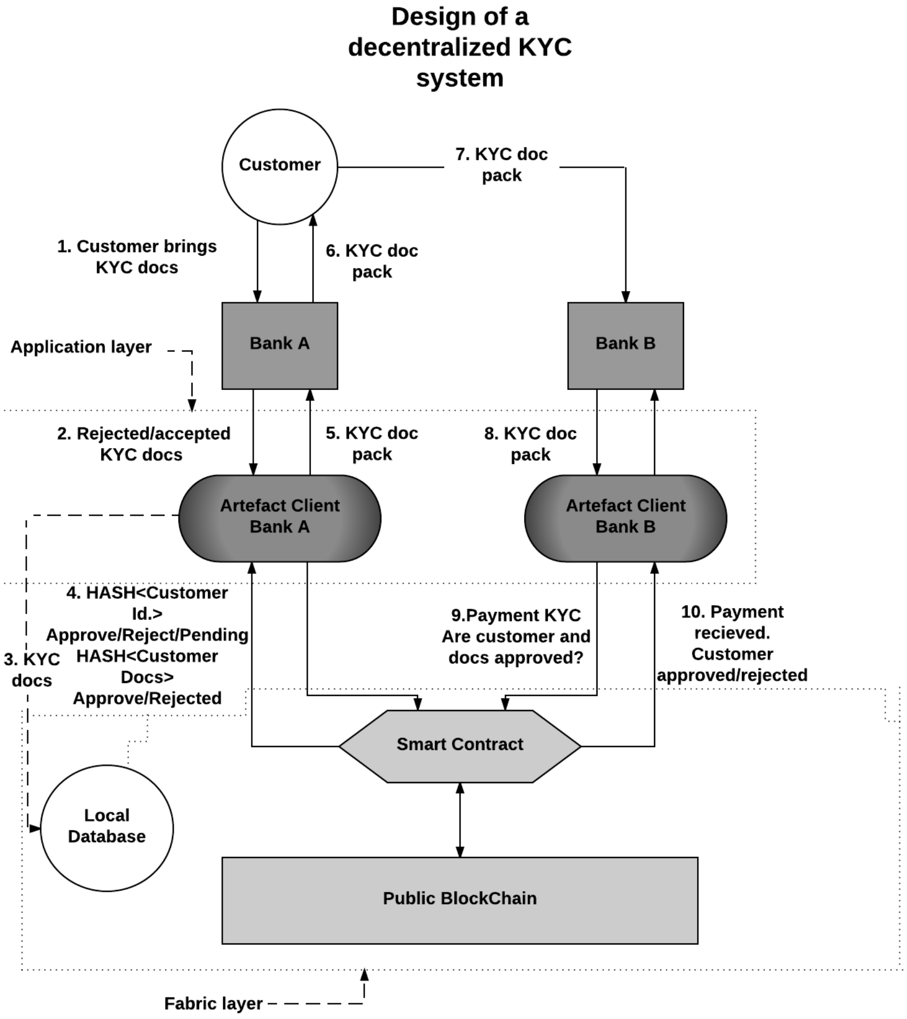

Design of a decentralized KYC system
Customer
7. KYC doc pack
1. Customer brings KYC docs
6. KYC doc pack
Bank A
Bank B
Application layer
2. Rejected/accepted KYC docs
5. KYC doc pack
8. KYC doc pack
Artefact Client Bank A
Artefact Client Bank B
4. HASH<Customer Id.> Approve/Reject/Pending HASH<Customer Docs> Approve/Rejected
3. KYC docs
9.Payment KYC Are customer and docs approved?
10. Payment recieved. Customer approved/rejected
Smart Contract
Local Database
Public BlockChain
Fabric layer

Figure 5: Design of the distributed KYC system. The blockchain is public, the documents are only kept by the home bank and the regulator does not have privileged access.

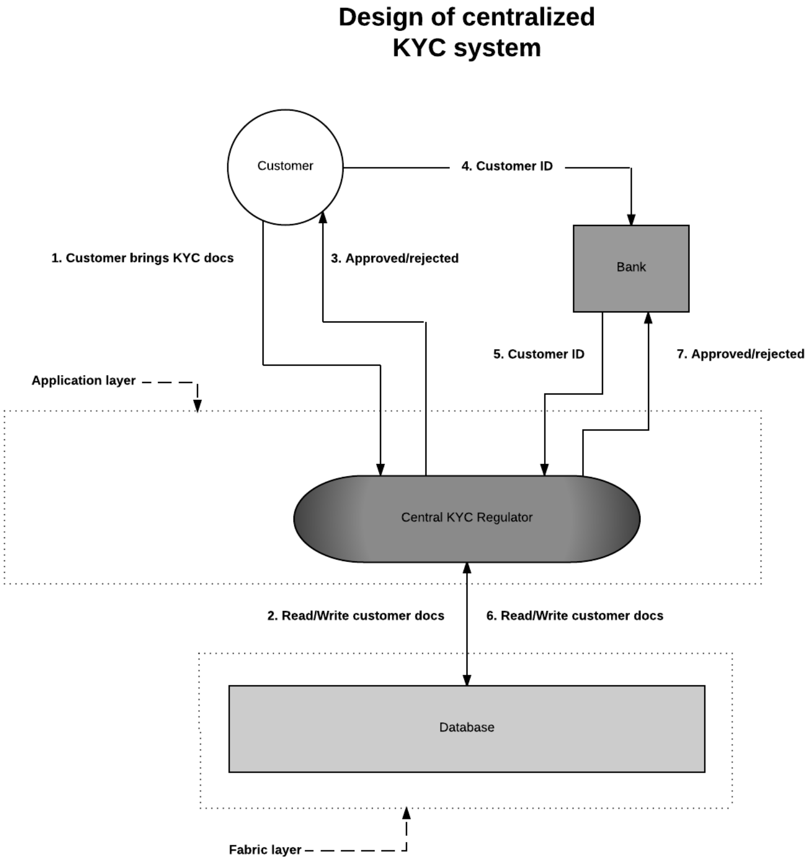


Figure 6: Design of the centralized KYC system