Università degli Studi di Milano - Bicocca
Scuola di Scienze
Dipartimento di Informatica, Sistemistica e Comunicazione
Corso di Laurea in Informatica

Distributed Ledger Technology: analysis and design of a DLT solution for a credits interchange sysstem

Relatore: Prof. Claudio Zandron

Co-relatore: Dott. Riccardo Mazzei

Relazione della prova finale di:
Nassim Habbash
Matricola 808292

Anno Accademico 2017-2018

To x
To y
To z

Abstract

Bitcoin and blockchains have been a disrupting force in the financial market. While these have a certain standing of their own - mainly through cryptocurrencies - the underlying technology, Distributed Ledger Technology (hereon DLT), is shaping up to transform the financial services sector. In most financial contexts, financial infrastructures are trusted by the counterparts of a transaction with maintaining, updating and preserving the integrity of the data in a central ledger, in addition to managing certain risks on behalf of the counterparts. This centralized model carries with it numerous inefficiencies, such as, to name a few: (i) transactional frictions, (ii) maintenance of record-keeping between different infrastructures (e.g. banks), (iii) infrastructure complexity, (iv) end-to-end processing slowness (that is, the speed in the obtaining and availability of assets and funds), (v) management of operational and financial risks. Thus, the financial and technology (FinTech) industry has been exploring different ways of leveraging the DLT to reduce the overhead and cost springing from the centralized model, increasing the drive for the research and development of this technology.

Objective

The aim of this project was to provide an analysis to Distributed Ledger Technologies, with the objective of identifying their risks, opportunities and implementation viability on different scales. The research work starts from a brief introduction to the technology and how the industry has been adapting to it, with statistical and analytical evidence. The paper then develops on the comparative analysis between the main competing technologies.

In accordance to the research findings, the project also aimed to provide a proof-of-concept architectural design of a solution for a credits interchange system using the most suited DLT implementation between the analyzed ones.

Contents

Chapter 1

Introduction

1.1 Emergence of the Distributed Ledger model

Ledgers have been a central element of commerce since ancient times, and are be used to record a variety of informations, from assets to property, but most importantly how these change hands, that is, transactions.

Ledgers have values as *archives*, in other words ledgers have value in their capability of being consulted to check, verify and manage the recorded transactions.

The medium on which transactions have been stored may have changed from clay tablets to hardware storage, but in all this time there haven't been notable innovations to the underlying architecture of the system. Each financial entity (i.e. banks, governments, financial institutions) manages its own ledgers, with its own technologies and implementative properties, based on their vision, necessities and customers (the would-be *counterparts* in the transaction), and in turn, the counterparts keep recorded their own views of the transactions.

This duplication of information between each party partecipating in the transaction drives a need for costly matching between each copy of the information, reconciliation and error fixing. The plurality of technology platforms upon which financial entities rely adds to that, creating more complexity and operational risks, some of which potentially systematic.

Think about the need for a party to transfer an asset, be it cash or a stock, to another. The transaction itself can be executed in microseconds, but the settlement - the ownership transfer of the asset - usually takes more time, from days to weeks. This length is due to different reasons: the parties don't have access to each other's ledgers, and can't automatically verify that the assets about to be transferred are in fact owned and not fake. So a number of intermediaries are needed as guarantors of the assets and to verify the transaction. A number of step has to be added just for this trusting mechanism, and in addition, the difference between infrastructures and technologies of every counterpart acting in the transaction can be such that there's always a need for reconciliation process between each party (ie adjusting each ledger to the transaction), increasing costs and length of the operations.

Centralized infrastructures were until recently an unavoidable model, as there were few ways to consolidate technologies without effectively consolidating the financial entities themselves. The industry has been moving towards the standardization and sharing

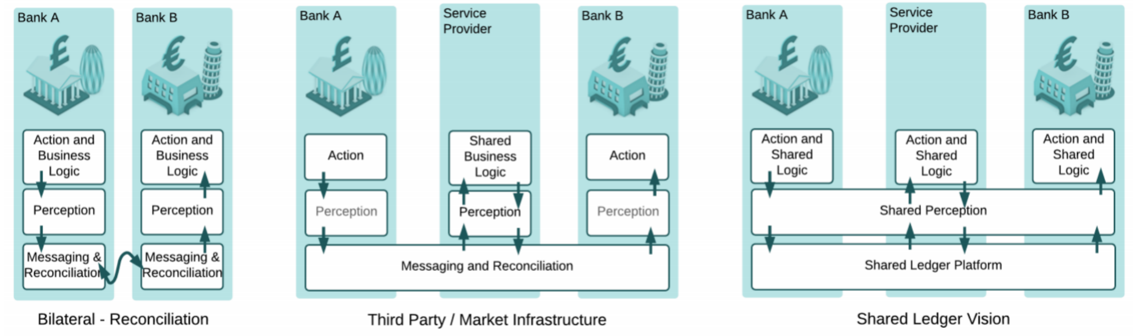

Figure 1.1: Comparison between architectures, Richard Gendal Brown, James Carlyle, Ian Grigg, Mike Hearn, “Corda: An Introduction”.

of data and some of the business-logic behind the architectures through the delegation of some part of the process to third-parties, but these steps are still lagging behind the evolution of the technology.

The term Distributed Ledger Technology refers to the processes, protocols and technologies that enable nodes in a distributed network to share data (propose, validate and record) between multiple synchronized data stores, collectively maintaned. The emergence of this technologies has had a stimulating effect in the FinTech industry, prompting the reconsideration of the entities needed in a financial transaction, how should trust be enstablished, the representation of the transaction, the securing of data, and many more. Whether DLT is not the answer in every case, asking these questions alone can be a force to drive progress forward.

1.2 Brief history

In 2008, a white paper (Satoshi Nakamoto, “Bitcoin: A Peer-to-Peer Electronic Cash System”) written by an as yet unidentified person using the pseudonym Satoshi Nakamoto, outlined a novel approach of transferring cash from one party to another without the need for a known and trusted third-party in a P2P manner, arguing, amongst other things, to have solved the issue of double-spend for digitalized currencies. Double-spending is the idea that a digital currency can be spent in twice - a unique problem of digital currencies, because digital information can be reproduced relatively easily, unlike physical currency.

The technology outlined in the paper was named Blockchain, referring to the way of organizing data and transactions. Bitcoin has soared in terms of popularity and value through its cryptocurrency market, that is, though, just one element in its whole architecture.

The effort of the industry since the introduction of blockchains has been directed to exploring different ways of leveraging this technology beyond Bitcoin, focusing on the core architecture of distributed record management. This use has gathered significant attention, reflecting the financial industry traditional reliance on multiple ledgers to maintain transactions. The use of DLT would be particularily effective for payment, clearing and settling activities because of the potential for simplification of the settling and reconciliation process between the parties involved.

Some of the resulting implementations of DLTs have been on a steady rise, such as Ethereum, that similarily to Bitcoin has seen a steep rise to the value of its cryptocurrency, Ether, and unlike its predecessor, offers a more malleable enviroment (which is the main reason Vitalik Buterinis created it), allowing for the transfer and recording of other assets like loans or contracts. Other rising implementations include R3 Corda, IBM Hyperledger Fabric and Digital Asset Platform.

As the research and development of the technology progresses, real-world applications have highlighted some of the challenges associated with these use-cases, including the need for safe, secure and scalable systems.

As of 2018, the DLTs impact in the financial sector seems still circumvented. According to a recent survey from Keith Hale, Sern Tham, *A Turning Point for the Global Asset Management Industry - The Multifonds Every Fund Survey 2017*, big data analytics, AI and robo-advice are currently higher up in the Asset Managers' agenda than DLT, explaining that it is "probably down to the fact that these new technologies are relatively easier to implement than more revolutionary DLT concepts". Despite strong progress in the research, it would seem that in the near-to-medium term many of the benefits and efficiency gains of DLT are likely to be reaped by start-ups and financial institutions in the developing countries, such as ABRA, *Abra*, ripple, *RippleNet*, Chain, *Sequence*, BitPesa, *BitPesa* and ShoCard, *ShoCard*.

What is a blockchain?

The term blockchain refers to the most well-known arrangement of DLTs (due to Bitcoin's fame), and refers to a distributed ledger architecture where the data is stored in entities called transaction blocks, linked with each other through chained encryption. The blockchain itself is the data structure formed by these linked blocks. Blockchains make use of algorithmic methods and encryption to ensure immutability, security and truthfulness of the information.

New additions are initiated by one of the nodes that creates a new block of data, containing the encrypted transactions. Information about the block is then shared on the network, and all partecipants collectively try to determine the block's validity according to a pre-defined alghoritmic validation method (consensus). After the validation, all the partecipant can add the block to their copy of the ledger. With this mechanism, every change to the ledger is replicated across the entire network, and each node has a full, identical copy of the entire ledger at any point in time. As the chain grows and new blocks are added, earlier blocks cannot be altered.

The cryptocurrency aspect is what has made Bitcoin garn the most fame. The consensus process involves the generation of a reward to the node that validated the last block of the blockchain, that being the currency in itself.

1.3 Technical design elements

It is emphasized that DLT is not a single, well-defined technology, but as of today there is a plurality of blockchains and Distributed ledgers (henceforth DLs) in active development.

DLs can be designed in a number of ways pertaining to main idea behind them and the use-cases they're designed to respond to. Such arrangements usually involve several key technical design concepts that specify how the information has to be kept on the ledger and how the latter has to be updated. There usually are four core attributes of DLTs, these are:

1. The distributed nature of the ledger
2. The cryptographic mechanisms
3. The consensus mechanism
4. The network access permission level

These three elements play are fundamental in ensuring the DLs ability to store and exchange data across different, self-interested parties, without the need for a central record-keeper, without the need for trust amongst the concerned parties, as it is guaranteed by the system itself, and while assuring that no double-spending takes place.

1.3.1 Distributed nature of the ledger

In its simplest form, a distributed ledger is a data store held and updated by each participant (or node) in a network. The control over the ledger does not lie with any one entity, but with several, if not all the network partecipants. This sets the technology apart from cloud computing or data replication, which are commonly used as shared ledgers.

No single entity of the network can amend past data entries, and no single entity can approve new additions to the ledger, which have to go through a predefined consensus mechanism. At any point in time there exists only one version of the ledger, and each network partecipant owns a full and up-to-date copy of it. After validation the new transaction(s) are added to all the ledgers to ensure data consistency across the network.

This distributed nature of DLs allows the removal of a trusted central party, increasing speed and potentially removing friction costs and inefficiencies associated with the matching and reconiculiation processes. It also improves security, removing the single point of attack and single point of failure that is represented by the central trusted entity. To potentially gain control over the network, a malicious third party would have to gain control over 50%+1 nodes in the network.

Security risks aren't completely solved: the software layer built over the DL can become additional attack surfaces.

1.3.2 Cryptographic mechanisms

Cryptography is at the core of the DLT. Asymmetric cryptography plays an important role by identifying and authenticating partecipants, confirming data entries and facilitating ledger updates. Each data entry is hashed, producing the so-called digest, which looks random and unrelated to the original input, but is in fact deterministic, meaning that from one original input there's only one hash possible. The data is hence hidden to anyone that is not intended to look at the data. Digital signatures, which are a common and robust method used in a wide array of application are used as a means of authentication. Each network partecipat has a private key, that is used for signing digital messages and only known to the key proprietary, and a public key, which is public knowledge and used for validating the identity of the sender of the original message. Partecipants proposing changes will authenticate themselves using digital signatures, and the validators will use cryptographic tools to verify whether the partecipant has the proper credentials, and so on. The validators can be either a counterpart, a third party, or the whole network depending on the type of DL and operation the change refers to.

In the blockchain subset of DLs in particular, encryption plays a fundamental role, as they're essential in the chain encryption mechanism between the blocks that make up the blockchain itself.

1.3.3 Consensus mechanism

The purpose of the consensus mechanism is to verify that the information being added to the DL is legitimate. It is fundamental in handling conflicts between multiple simultaneous competing entries (ie double spending), or take-overs by bad actors in the network. It's a derivative property of the distributed nature of the ledger, and it requires partecipants in the network to reach a consensus over the information being added. There exist many different consensus algorithms and mechanisms, with different purposes, advantages and disadvantages.

Consensus usually involves two steps:

1. Validation, where each validator involved identifies that the state change is consistent according to the rules of the ledger. This operation may rely on records of previous states or a last agreed state.

2. Agreement, where each validator agrees to the state changes to the ledger. This step involves the mechanisms to resolve eventual conflicts and ensuring that valid changes are made only once, thus ensuring that the whole network is synchronized.

The Bitcoin blockchain uses Proof-of-Work (PoF) to enstablish consensus. To add a new block to the blockchain, a node has to provide a proof of work. This is a computationally taxing problem, but easy to verify, and is solved by brute-forcing cyptographic hashing algorithms until the correct string that satisfies certain conditions is generated. This process is called "mining". Each miner that produces a valid PoF is then rewarded Bitcoins, which serves as an economic incentive to maintain system integrity.

The Ethereum blockchain uses Proof-of-Stake (PoF). Its process is quite different from the PoF of Bitcoin, as there's no mathematical problem to solve, but instead, the creator of the new block is chosen in a deterministic way based on their stake, that is, how many coins or tokens they possess. A key advantage to this approach is the energy efficiency. The Bitcoin network, for example, requires an annual energy consumption comparable to that of Columbia (56.6 TWh annually). Thus PoS systems are well suited to platforms where there is a static coin supply, without inflation from block rewards. The rewards consist only in the transaction fees.

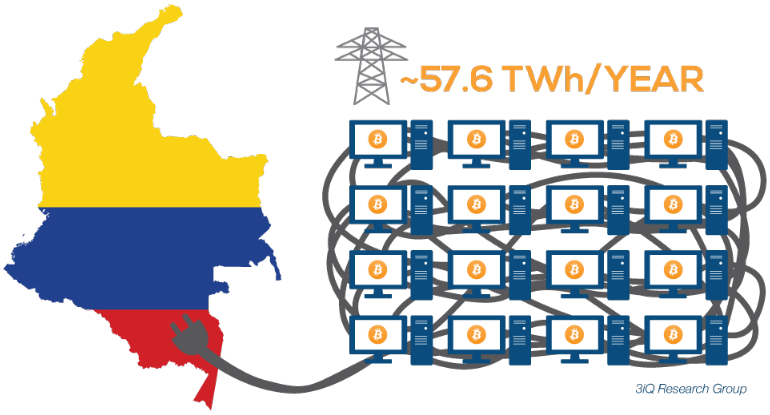

Figure 1.2: Data as at April 3rd, 2018. Retrieved from https://digiconomist.net/bitcoin-energy-consumption

The Corda distributed ledger, utilizes a unique "pluggable" consensus service, due to its peculiar distributed ledger architecture. It divides consensus in two types, validity consensus and uniqueness consensus. Given a proposed transaction, the first type of consensus checks whether the two parties satisfy a set of validity conditions, going back through all the parties' transaction histories, while the latter has the purpose to avoid double spend, and the consensus is provided by a Corda network service called "Notary", by attesting that, for a given transaction, there hasn't been proposed another competing one.

1.3.4 Network access permission level

The partecipation in the DL network can be open (permissionless) or permissioned. Bitcoin and Ethereum are the most prominent examples of completely permissionless blockchain, where partecipants can join or leave the network at will. This plays as one of their strengths, as a large, open permissionless system with a large number of nodes incentivized to validate new changes to the ledger and accurately and enstablishing a consensus is directly related to its network security (1.3.3).

In permissioned DLs its members are pre-selected by somone - an owner, or a network service - who controls network access and sets the rule of the ledger. The regulations of network access usually permits the use of a non-computationally expensive consensus mechanism, as there is no requirements for any trust between partecipants. This, however, means there's now a centralized trust entity playing a coordinating role and bearing the responsability over the trusting mechanism. In permissioned DLs it's possible to have different degrees of transparency over the ledger, and faster transaction processing (thanks to the lighter consensus algorithm) allows for higher transaction volumes.

The identity verification needed for the access solves some problems with governments and regulators with concerns about the identity verification and legal ownerships clarifications.

Permissionless DLs have open access to the network, so anyone can join and leave as they wish. There's no central owner or administrator, the ledger is wholly transparent and the security is enstablished by having a large scale network. It is required to have a complex consensus algorithm to guarantee the integrity of the information, and there are some legal concern over lack of ownership, as no legal entity owns or controls the ledger.

Some industry players make distinctions between public/private, in term of access, and permissioned/permissionless, in term of roles in the network. For example, Ripple, has a permissioned ledger, but the data is validated by all partecipants, therefore being a public, permissioned ledger. Corda, on the other hand, has a permissioned ledger, but the data is validated only by a set of partecipants (those which the data concerns), hence being a private, permissioned ledger.

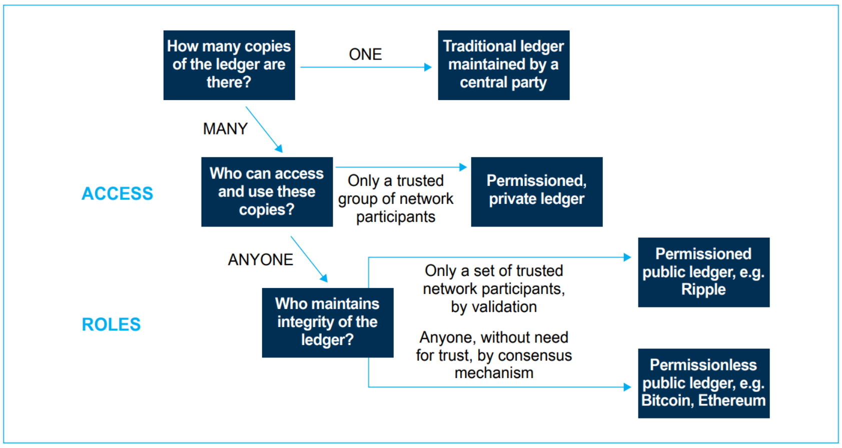
How many copies of the ledger are there?
ONE
Traditional ledger maintained by a central party
MANY
ACCESS
Who can access and use these copies?
Only a trusted group of network participants
Permissioned, private ledger
ANYONE
ROLES
Who maintains integrity of the ledger?
Only a set of trusted network participants, by validation
Permissioned public ledger, e.g. Ripple
Anyone, without need for trust, by consensus mechanism
Permissionless public ledger, e.g. Bitcoin, Ethereum

Figure 1.3: Network Access Ledger Taxonomy, Dave Birch, "Distributed Ledger Technology: beyond block chain".

1.3.5 Roles

Nodes in the network may play a variety of roles, depending on the partecipants intentions or technical arrangement of the DL. The Committee on Payment and Market Infrastructures of the Bank for Internation Settlements proposed a generalized framework, with the following different, non-exclusive roles for a node:

1. System Administrator: node controlling access to the system and provides dispute resolution and notary services. This role is not required in permissionless DLs.

2. Asset Issuer: node enabled to issue assets. In the Bitcoin blockchain, there's no entity playing this role as the system creates assets (Bitcoins) by itself, according to its rules.

3. Proposer: node enabled to propose ledger updates.

4. Auditor: node enabled to view the ledger, but not to make updates. Can be used by regulators or supervisors.

5. Validator: node enabled to validate requests for addition of transactions in the ledger. This role is performed by the consensus mechanism in permissionless DLs.

1.3.6 Process flow

Bibliography

[1] ABRA. *Abra.* URL: `https://www.abra.com/`.

[2] BitPesa. *BitPesa.* URL: `https://www.bitpesa.co/`.

[3] Chain. *Sequence.* URL: `https://chain.com/`.

[4] Dave Birch. "Distributed Ledger Technology: beyond block chain". In: ().

[5] Keith Hale, Sern Tham. *A Turning Point for the Global Asset Management Industry - The Multifonds Every Fund Survey 2017.* 2017.

[6] Richard Gendal Brown, James Carlyle, Ian Grigg, Mike Hearn. "Corda: An Introduction". In: (2016). URL: `https://docs.corda.net/_static/corda-introductory-whitepaper.pdf`.

[7] ripple. *RippleNet.* URL: `https://ripple.com/`.

[8] Satoshi Nakamoto. "Bitcoin: A Peer-to-Peer Electronic Cash System". In: (2008). URL: `https://bitcoin.org/bitcoin.pdf`.

[9] ShoCard. *ShoCard.* URL: `https://shocard.com/`.