Introduction

A blockchain is a form of network-distributed ledger, whose users play an active role in keeping it constantly updated. The initial concept was designed in 2008 and implemented in 2009 as the core protocol of the digital currency Bitcoin. This first blockchain was conceived with the purpose of allowing peer-to-peer transactions through Bitcoin, and it has since been a source of inspiration for thousands of different developers.

The term "blockchain 2.0" refers to all the most recent evolutions of the blockchain, whose potential applications go far beyond exchanging value without an intermediary (Swan, 2015). Its benefits might include advanced security (Dorri et al., 2016), data transparency (Pilkington, 2016), digital intelligence (Swan, 2015), disintermediation and many others (Tasca et al. 2017).

Based on these benefits, according to a recent report from PwC, "blockchain could become a force anywhere trading occurs, trust is at a premium, and people need protection from identity theft" (Plansky et al., 2016). Such a potential is making blockchain the most promising technology in the digital arena together with Artificial Intelligence, as recognized by important institutions and market analysts (World Economic Forum, 2017; Gartner Group 2016).

Prominent thought leaders are endorsing this technology every day. Robert Greifeld, former Nasdaq CEO stated that blockchain is the «biggest opportunity set we can think of over the next decade». Don Tapscott, one of the most influential opinion makers in the digital landscapes, said of blockchain: "I've never seen a technology that I thought had a greater potential for humanity." Joi Ito, Director of the MIT Media Lab, recently wrote that "the blockchain will do to the financial system what the internet did to media."

The increasing enthusiasm of the business community around blockchain technologies is also powered by several concurring trends.

First, looking to the native application field of blockchain technologies, the global market of cryptocurrencies has grown significantly during the past few years and has exceeded $295 billion in April 2018, starting out at $18 billion at the beginning of 2017 and hitting $795 billion in January 2018.

Second, both the top ICT players and the largest venture capitalists are heavily investing in new companies focused on blockchain technologies, applications and standards. According to Crunchbase's data, during 2017 the worldwide venture investment in blockchain and blockchain-related startups hitted $1 billion starting from less than $600 millions in 2016.

Third, looking to new application fields of the blockchain, several big names outside financial services, such as Walmart and Maersk, have started to launch implementation initiatives aimed at testing the benefits of distributed ledger technologies (Fortune, 2017).

Lastly, several companies, research institutions and industry consortia are joining forces to create blockchain standards, platforms and applications. Examples of broad networking initiatives aimed at advancing blockchain technologies for either cross-industry applications or industry-specific applications are, respectively, Hyperledger and R3.

Unfortunately, the combined effect of these trends is leading to a hype effect around blockchain (Morini, 2016; Notheisen, Hawlitschek and Weinhardt, 2017). While it is commonly accepted that blockchain could lead to radical changes in many industries (Mattila, 2016), with a potential impact on the whole economy (Swan, 2015; Tapscott and Tapscott, 2016), several authors focus on the medium-to-long time needed in order to actually experience some transformational impacts of this technology. This is mainly due to the foundational nature of blockchain (Iansiti and Lakhani, 2017):

> "It has the potential to create new foundations for our economic and social systems. But while the impact will be enormous, it will take decades for blockchain to seep into our economic and social infrastructure. […] Many barriers—technological, governance, organizational, and even societal—will have to fall."

Consistently, most of the effort expended by the academic world in the last 5 years have been devoted to solving the challenges that are slowing down the potential disruption led by blockchain and distributed ledger tech-

nologies, with a main focus on Bitcoin and other cryptocurrency applications. Very few sources have focused their attention on a comprehensive assessment of the current application landscape of blockchain technologies (Salviotti, De Rossi and Abbatemarco, 2017).

As a consequence, business leaders and practitioners are still trying to address several unanswered questions:

- Where should I start my journey into this world?
- Who owns the blockchain in current business implementations?
- What are the main technical features of blockchain platforms currently implemented?
- What are the main business applications of blockchain, other than cryptocurrencies?
- What are the available options to build a solid blockchain strategy for my company?

According to Bill Gates "blockchain is a technology tour de force." Every blockchain journey should start from a sound understanding of the technical pillars that make this foundational technology capable of creating digital assets that are not duplicable. Chapter 1 is devoted to this point.

In Chapter 2 we start analyzing blockchain from different dimensions, analyzing aspects such as type of ownership, protocol independency and decentralized consensus mechanisms.

After gaining an understanding of the key features and the functioning mechanisms of our "vehicle" we can start dreaming about cool and unexplored destinations. The aim of Chapter 3 is to provide a list of the most valuable business applications of blockchain: cryptocurrencies, data certification , digital advertising, digital identity, digital voting and governance, energy management and distribution, financial payments platforms, gaming, IoT platforms, P2P content distribution, P2P resource distribution, prediction markets, and smart contract platforms. For each application, the chapter provides examples of relevant projects and use cases collected from the field by the authors during their researches at SDA Bocconi's DEVO Lab.

In Chapter 4 we show an overview of the current public blockchain landscape. The chapter focuses on an original market analysis conducted by the authors and aimed at collecting examples of the most interesting running blockchain, divided by the abovementioned application areas.

The first four chapters should provide the reader with all the definitions and technical concepts required to understand blockchain. By the time we reach Chapter 5, the reader will be ready to plan his full journey into this world; working on the different layers of the blockchain architecture, we present a continuum of four architectural options useful to shape the design of a company's blockchain strategy.

Are you ready to start your journey?

1 What is the Blockchain?

This chapter aims at providing a clear understanding of what this technology is, how it works and what the key features that can be used to analyze the different types of blockchains are. To achieve this goal, after a short history of Bitcoin, the chapter focuses on the main technical building blocks that characterize a blockchain: distributed computation, advanced cryptography, data architecture and consensus mechanisms.

1. Blockchain: the background history

The history of blockchain is closely related to the programmer (or group of programmers) going by the pseudonym of Satoshi Nakamoto, founding figure of Bitcoin and consequently of the blockchain technology. It is indeed quite difficult to talk about blockchain without naming Bitcoin: Bitcoin was the blockchain original protocol, and is still the most famous of its technological applications.

As highlighted by Nakamoto himself, it is clear that much of Bitcoin's success depended on the maturity of its underlying components, such as cryptography, peer-to-peer (P2P) networks, redundant databases, and digital currencies, to name but a few. In fact, the idea of an electronic or digital cash/payment system is not original to Bitcoin, and dates back almost 40 years. The concept was originally elaborated by the cypherpunk movement, a group of people that in the early 1990's established a regular mailing list to discuss topics ranging from mathematics, cryptography, and computer science to political and philosophical visions.

The journey from here to Bitcoin, and therefore to the blockchain technology, has been marked by at least six key milestones.

1982: Ecash (David Chaum)

The first proposal to create digital cash dates as far back as the early 1980s. In 1982, David Chaum, an American computer scientist and cryptographer, proposed an innovative scheme to build an untraceable digital currency, thanks to the use of so-called blind signatures. The idea was to enable regular banks to issue digital money, in the form of signed (i.e. recorded into a ledger) random serial numbers; through the "signature," the bank could certify the authenticity of the serial numbers after their release to the public.

The currency thus created could have been exchanged between users in a completely anonymous way, given that the bank could only certify the authenticity of the serial number and not its movements.

The main limitation of Chaum's model was that the currency was still based on a central system – a bank – which needed to keep track of all the serial numbers used to issue new currency. Through his company Digicash, Chaum managed to actually deploy Ecash as a micropayment system from 1995 to 1998, before going bankrupt. During this period, one of the biggest critical issues related to digital money also emerged: the double spending[1] problem.

1997: Hashcash (Adam Back)

In 1997, Adam Back – a British cryptographer and crypto-hacker – created a new algorithm, called Hashcash, to thwart unwanted e-mails. The idea behind Hashcash was simple: to integrate into each e-mail the solution to an easy to verify, but extremely expensive and difficult to compute computational puzzle. A single user would not have even noticed the extra computational effort required to send an e-mail, while on the contrary an operator sending a high number of spam e-mails would have been discour-

[1] In short, digital cash is nothing more than a series of numbers, and therefore it results very easy to duplicate it. If the validating entity cannot track all the transactions, history, as in the Ecash case, then a user can make a copy of its digital money and send it to a merchant or another party while retaining the original.

aged to do so as the time and resources required to run the spam campaign would have increased substantially.

1998: B-money (Wei Dai)

In 1998, it was the turn of Wei Dai, a computer engineer often described as intensely private, to introduce a new currency called B-Money. Dai associated the Hashcash algorithm to Ecash's digital money concept – naming this combination Proof-of-Work. The Proof-of-Work system solved the problem of double spending: issuing money became, just as for sending e-mails, an extremely complex and expensive process.

In B-money, cash was transferred by broadcasting the transactions to all participants, each of whom keep accounts of all others. However, a new weakness arose: potentially, an adversary with higher computational power than the whole network of B-money users could have generated new money in unlimited amounts, without allowing the network to respond. Despite this relevant issue, B-money can still be considered as the first real attempt to translate physical scarcity in the digital world.

1998: Bit Gold (Nick Szabo)

Immediately after Wei Dai's release, another computer scientist, Nick Szabo, introduced the concept of Bit Gold – an improved version of B-money.

Szabo, who in the following years would become the first to develop and disseminate the concept of smart contracts, managed to integrate a solution for the adversary attack. In Bit Gold, the network could dynamically respond to the attack of an adversary and be much more secure.

1999: Auditable, anonymous Ecash (Tomas Sander, Amnon Ta-Shma)

In 1999, Tomas Sander and Amnon Ta-Shma introduced an updated version of Chaum's Ecash project that used an innovative cryptographic scheme to represent coins and possessions: the Merkle tree. This scheme allowed users to be fully anonymous, albeit at a small cost in terms of required computational power. As for the original Ecash system, the main problem of this solution was that it needed a central bank to keep records of all the serial numbers used.

2008: Bitcoin (Satoshi Nakamoto)

Satoshi Nakamoto leveraged all the previously mentioned technologies in order to implement a P2P system for exchanging money in the form of digital tokens (i.e. bitcoins). Moreover, inspired by Sander and Ta-Shma's modified version of Ecash, Nakamoto equipped Bitcoin with a new database structure which allowed the entire transaction history to be kept in a completely distributed network. No single group of individuals, including governments, banks and corporations, controls Bitcoin because all the peers on the network operate as equal actors.

It is the combination of all the components listed so far, whose detailed description will be provided in the following paragraphs, that gave life to the blockchain technology. Even if Satoshi Nakamoto is commonly recognized as the blockchain inventor, the world "blockchain" did not even exist back in 2008, and it was not mentioned in the original Bitcoin's White Paper. It only appeared later, embedded in comments of Bitcoin's source code, as a "chain of blocks."

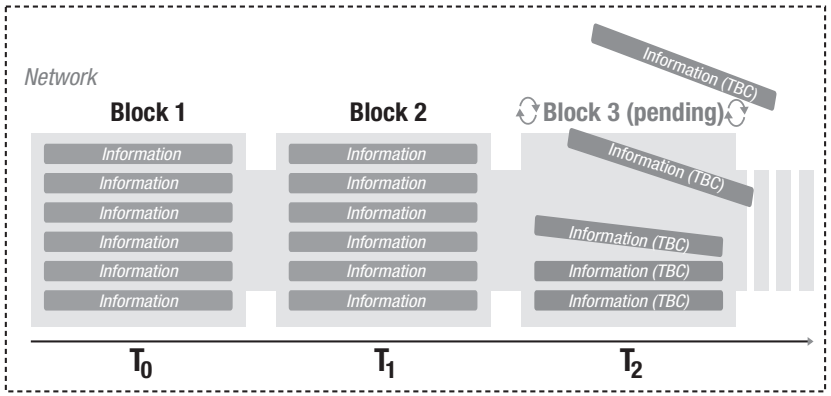
2. The Blockchain: definition and technical pillars

A blockchain is a decentralized database structured in blocks, each one containing a certain amount of information and distributed through a chain (i.e. a ledger) over a network. Therefore, it is a digital way to store any kind of data, be it a token of value or a crypto money balance, through a network (for example, the Internet).

Usually, stored data are semi-public: every user (also called "a node") in a blockchain infrastructure can verify if another user stored an information, but only that specific user can unlock the content of the information, since he is the only one who holds the keys to access that data.

Data stored in a blockchain cannot be lost. They are there forever, replicated as many times as the number of nodes in the network. Moreover, the blockchain does not simply store the last state of the data, but rather the whole history of all its previous states, so that every node can check the integrity and the correctness of the final state by repeating each intermediate step since the beginning.

Figure 1 Blockchain infrastructure



Source: DEVO Lab, 2017

The building blocks on which the blockchain is based are:

- Distributed network;
- Advanced cryptography;
- Data architecture;
- Decentralized consensus.

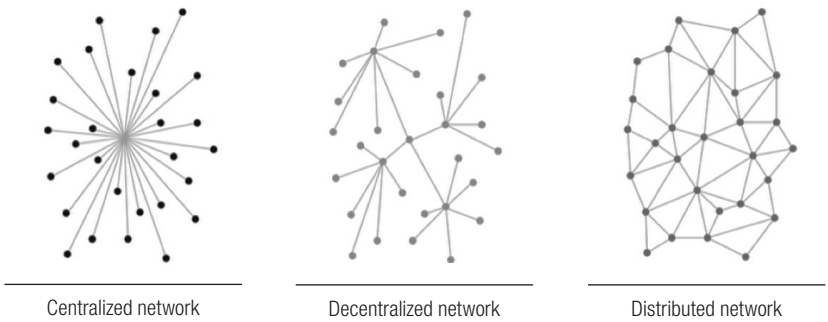
Distributed network

A key component of blockchain is its P2P or distributed network topology which eliminates the need for a trusted third party or intermediary service to perform certain functions.

The type of network in which the database will be shared heavily influences the shape it can assume. As we can see in Figure 2, it is possible to identify three main network paradigms:

- *Centralized*. In a centralized ecosystem, nodes are arranged in a client-server relationship, where the server acts as the single governing authority responsible for each and every operation. Everyone who wants to submit a record must pass through a central entity. Centralized systems directly control the flow of information managed by the individual units from a single central point. All the individuals are dependent on the central power, both in terms of sending and receiving information;

Figure 2 Centralized, decentralized, and distributed networks



Source: Baran, 1964

- *Decentralized*. In a decentralized system, data are spread across multiple nodes in a network. However, this model still retains a central authority that controls and governs processing over all other nodes in the system. Therefore, the system's nature may still be viewed as essentially centralized. An example of a decentralized network is that of a bank's electronic system, where the bank remains the central authority that controls and governs the entire system.
- *Distributed*. In this type of network, data are constantly shared and synchronized across nodes, even if they are spread across multiple sites, institutions or geographies. The owners of each node can access the records shared across that network and replicate them. Any change or addition made to the network is automatically mirrored in all its copies. What is actually "distributed" in this type of network is the responsibility of managing the overall infrastructure: the absence of central accountability forces each user to rely on a general consensus of the peers involved in the ledger's management. Thus, distributed networks are not defined in terms of how or where the information they contain is stored; instead, they are defined by the network's consensus process – the process peers use to reach consensus – which is always shared within the whole ledger.

Blockchain networks own some typical characteristics of distributed systems, highlighted for example in Babaoglu and Marzullo (1993), Chandra and Toueg (1996) and Tenenbaum and Van Steen (2007), such as:

- "Full" nodes: each node has full knowledge of all the information present in the system;
- Structure neither fixed nor precisely determined: latency, type and size of the network can also vary during construction;
- Lack of single point of failure: a single node can fail without compromising the network operativity in any way.

A blockchain network results in fact composed of many systems (be they servers, smartphones, computers, etc.), geographically scattered, all connected through a single communication system in which various types of messages can be exchanged. In the most common case in everyday practice, the communication system is represented by the Internet.

For example, in the Bitcoin blockchain, each user must download all the transactions ever recorded on the blockchain to enter the network. After this step, each node can run independently, processing incoming transactions and propagating them further: there is no need for a central node processing and distributing data. Moreover, each node can contribute to reach the consensus.

Advanced cryptography

Data saved on a blockchain are never immediately decipherable: all the contents are encrypted through various advanced cryptographic techniques, so as to guarantee their fruition to the legitimate owners only. For example, transactions (which are basically data transfers) are digitally certified by the sender through the use of cryptographic signatures, similarly to how mails in the past were marked with an official sealing wax to verify the authenticity of the document and the identity of its author.

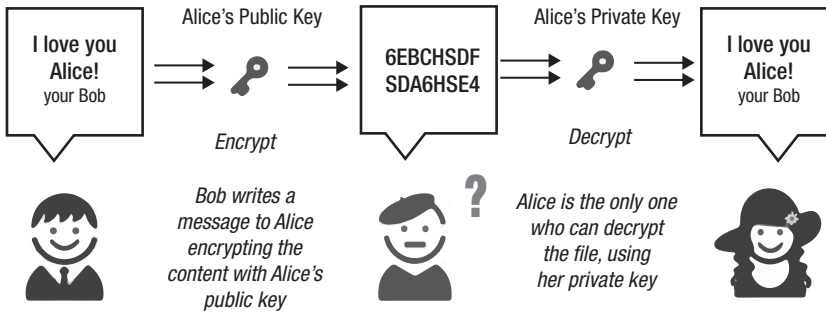
Furthermore, cryptography also enables some of the mechanisms at the heart of the blockchain data architecture. There are two main cryptographic tools used in a blockchain: public-key cryptography and hash functions.

Public–key cryptography

Also known as asymmetric cryptography, public-key cryptography is a branch of cryptography based on the use of a pair of numbers bounded together by a complex and non-bi-univocal mathematical relationship. Such numbers, once paired together, are called keys.

The first techniques of public-key cryptography were theorized by the cryptographers Ellis (1970) and Cocks (1973), but were kept confidential

Figure 3 Public-key cryptography



Source: DEVO Lab, 2017

by the British government and hidden from the academic community, owing to their possible exploitation from a military point of view. The first public works on the matter therefore go back to the studies of Diffie and Hellman (1976) and Rivest, Shamir and Adleman (1978).

As previously mentioned, the keys used in this encryption system are linked to each other by a mathematical function. The keys are generated in such a way that from the first of the two, called private, it is possible to derive the second, called public, but it is not possible to do the opposite. The keys thus obtained perform two important tasks:

- *Encrypting and decrypting a message*. A message encrypted using a public key can only be decrypted by the owner of the private key. The public key basically acts as a mailbox, while the private key acts as the key needed to access it;
- *Signing a message*. Given a message signed with a private key, anyone with the corresponding public key can check if the signature of the message owner is authentic or not. In this case, the use of the private key is comparable to that of a personal sealing wax, similar to those used in past centuries to verify the authenticity of a sender.

Therefore, public keys can be freely shared, providing users with an easy and convenient method for encrypting content and verifying digital signatures, while at the same time private keys should be kept secret, as they ensure that only their owners can decrypt content and create digital signatures.

For practical use, since public keys need to be shared often but are too big to be easily remembered, they are usually stored on digital certificates for secure transport (for instance, QR codes). Private keys, on the other hand, can be stored on dedicated software (so-called software wallets) or hardware (hardware wallets).

Hash functions

Cryptographic hash functions refer to a set of techniques where an input of varying length and size is converted and encrypted into an output of fixed length and size.

Hash functions are extremely useful to determine if two pieces of data are identical. For example, in a blockchain they are used to establish whether or not two transactions correspond; this is achieved by converting them into fixed-length outputs called "hash values." Table 1 illustrates how data inputs are transformed by the cryptographic hash function into a fixed-length message digest or hash value.

Table 1 Converting data inputs into hash values – SHA256

Input	Function	Hash Values
Salviotti	SHA256	D6867D2257AFC1A9D47D0C70BBA953DAE5C07BBC457A6 FE83A2E8F4F3640704E
salviotti	SHA256	74AC3C5939F49122F508B72F08EF429D17093D9A960775A6 B21CFED5C4604C5
De Rossi	SHA256	93E7890E61E057AF1D6383813C65E0CED549072F0A4D774 B281E69409191A701
DeRossi	SHA256	FE6C45988AF91FF710242E0BA28B73A21C5B239698571D5 F5FFF60C16C1F5341
Abbatemarco	SHA256	B96DDA24C7F606626E2BCBD6EABB714E6D73D93DAC43FA9 A40993CB3F744C4A0
Abbatemarco1	SHA256	0957A7BB30EF3FB9F9ED0966A68C4B88DF80E7AA8C88E4 FA8 06ABB1B222D1961

Source: DEVO Lab, 2018

Hash functions can vary according to the length of the hash values they produce. Blockchain protocols typically use the Secure Hash Algorithm-1 (SHA-1) and the Secure Hash Algorithm-2 (SHA-2) family of algorithms designed

by the United States National Security Agency (NSA). Bitcoin employs the SHA-256 hash functions, which generates an almost-unique 256-bit value from an input text.

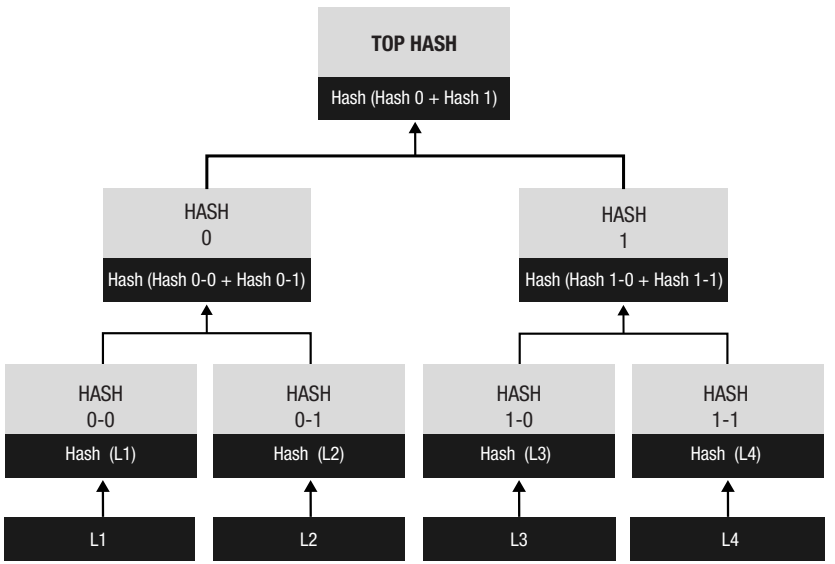
Data architecture

Blockchain deploys a specific data structure – called "Merkle tree" – to guarantee the integrity of the entire ledger. A Merkle tree is usually represented as an upside-down tree with:

- Single transactions at the bottom level;
- Hash values of the single transactions at the second level;
- New hashes, born form the combination of single-transactions hashes, at the third level;
- Single hash of the entire tree at the top level.

The final value – i.e. the top hash – represents the so-called "Merkle root" and provides proof of validity for all the transactions added to the tree (Pogson et al., 2017). An example of a Merkle tree is presented in Figure 4.

Figure 4 A Merkle tree



Source: Wikimedia Commons, 2012

Blockchain integrates this Merkle root within a data structure composed of different "blocks." In a blockchain, all the transactions are bundled together into a block which is linked with a cryptographic hash to the previous one. The cryptographic linkage between blocks results in the "tamper-proof" property of the ledger, because if a malicious actor tries to add, remove or change a transaction in any one block, this will affect all the blocks that follow. Figure 5 represents the actual structure of a blockchain.

Decentralized consensus

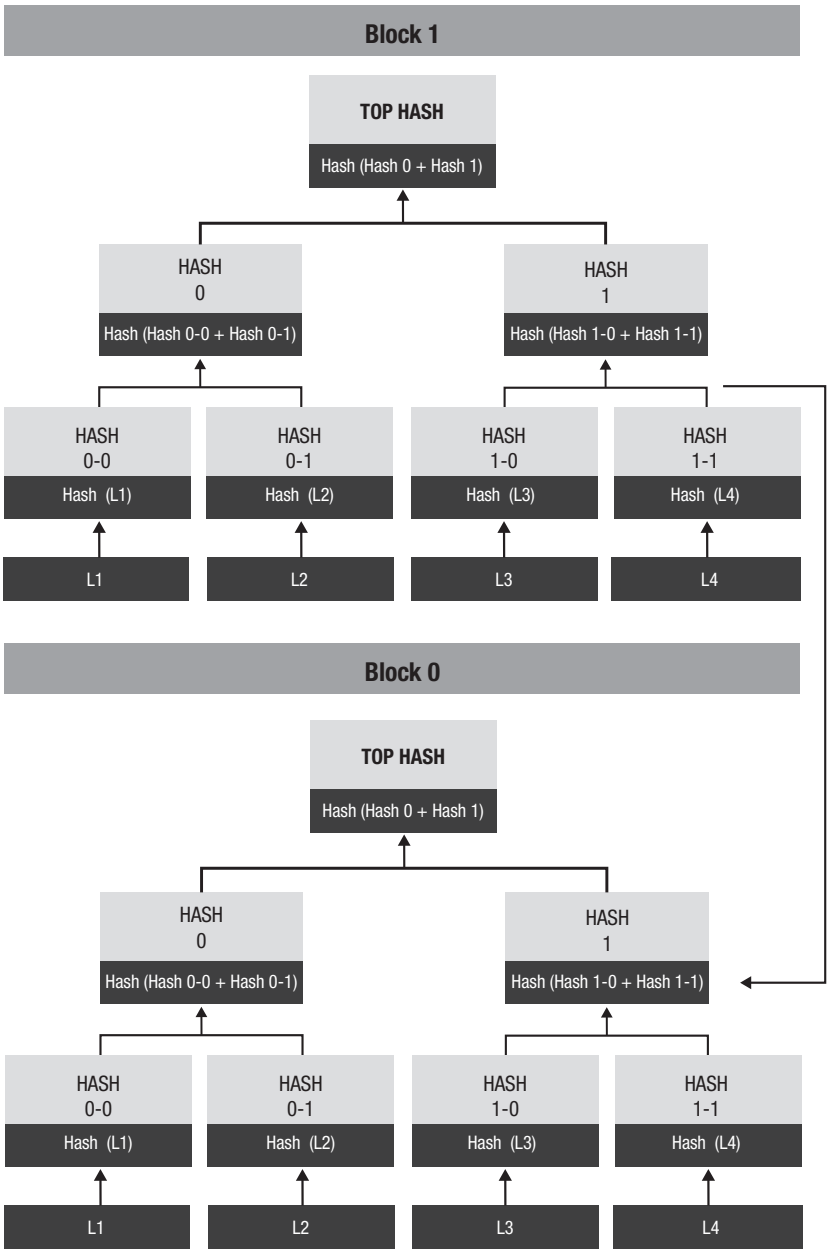
As mentioned before, a blockchain is a network-distributed ledger whose nodes continuously record information in "blocks," assembled in a unique "chain."

The consensus mechanism is the keystone of the entire blockchain technology, as it ensures that the information entered in the blocks is correct and consistent with the rules established in the protocol. It represents a solution, sought for over 20 years in the world of cryptography, to the so-called Byzantine Generals' Problem. To quote the original paper (Lampert et al., 1982):

> Reliable computer systems must handle malfunctioning components that give conflicting information to different parts of the system. This situation can be expressed abstractly in terms of a group of generals of the Byzantine army camped with their troops around an enemy city. Communicating only by messenger, the generals must agree upon a common battle plan. However, one or more of them may be traitors who will try to confuse the others. The problem is to find an algorithm to ensure that the loyal generals will reach agreement. It is shown that, using only oral messages, this problem is solvable if and only if more than two-thirds of the generals are loyal; so a single traitor can confound two loyal generals. With unforgeable written messages, the problem is solvable for any number of generals and possible traitors.

A consensus mechanism allows information to be shared between two nodes belonging to the network without necessarily having to transit for a central entity to validate the content first. While in a network where the interests of all nodes are aligned the problem of validation does not arise, the same cannot be said for one that potentially has a malevolent entity within it.

Figure 5 A schematic representation of a blockchain



Source: DEVO Lab, 2018

Although a theoretical solution to the problem of the Byzantine Generals had been already expressed by the authors of the original paper, from a practical point of view the implementation of the first functioning consensus mechanism arrived only in 2008, with Satoshi Nakamoto's White Paper "Bitcoin: A Peer-to-Peer Electronic Cash System". The mechanism became operational for the first time after the release of the Bitcoin protocol in January 2009. Even if slightly different, it kept the same name used by Wei Dai in its first implementation in B-money: Proof-of-Work (or PoW).

3. The blockchain tools

So far, we have illustrated in detail the technical aspects behind blockchain. What, however, will probably result still unclear is the kind of tools enabled by this technology.

In the early days of this technology, the only instrument empowered by blockchain seemed to be one: the digital token. A digital token is nothing more than a unique string of characters that, thanks to their uniqueness, becomes a de facto scarce digital asset, easy to represent on a ledger and that can be used to transfer value between two users who agree to confer it a certain value. According to many, this was (and still is) the only valid application offered by the blockchain technology.

Over time, however, developers realized how it was possible to conceive tools different from the simple token but working on the same logic: a set of instructions, nothing more than lines of computer code, but permanently inscribed on a ledger. The concept of smart contracts has been built exactly on this idea; the digital token became thus conceptually the first and one of the simplest forms of smart contract. We now move tho analize these tools in more detail.

Digital tokens

When present in a blockchain protocol, a digital token is represented as a unique string of code lines to which the community participating in the network attributes a certain value. Despite being considered one of the simplest forms of smart contract, the token still represents a crucial element in blockchain architecture. This is because it performs three fundamental tasks:

- It is used as the means of exchange within the protocol. Tokens cannot be represented exactly as physical nor digital objects; they can rather be conceived as entries on the protocol ledger. Owning a token therefore corresponds to owning the pair of keys able to decrypt that entry in the ledger, and exchanging it means to swap the keys that unlock it with another network user;
- It allows creators of blockchain content to be rewarded for their work. Unlike a common start-up, a blockchain protocol is a decentralized solution, and this means there is no way for its creators to maintain ownership of their project after having shared it with the reference network. However, should the developers be able to retain some of the tokens at the moment of the release to the public, there would be a twofold incentive: for developers, the ability to retain part of the value created; for investors, the certainty (theoretically, at least) that there is an incentive to create the best possible product (the more valid the project, the higher the value of the underlying token). This fundraising model, called ICO (Initial Coin Offering, inspired by that of Initial Public Offering in stock exchanges), has experienced a remarkable boom in 2017;
- It provides an economic incentive for the whole network to maintain the blockchain secure. As we will see later on in the paragraph dedicated to consensus mechanisms (Par. 4, Chapter 2), finding a consensus in a distributed network represent a time- and money-consuming job; were validating nodes not to be rewarded for their work through the issuing of new coins, they would have no reason to continue to keep the protocol safe, exposing it to potential hacker attacks.

Smart contracts

Smart contracts are a central component to next-generation blockchain platforms and are already becoming a cornerstone for enterprise blockchain applications.

Basically, a smart contract is a computer program code that is capable of facilitating, executing, and enforcing the negotiation or performance of an agreement (i.e. contract) using blockchain technology. Vitalik Buterin described smart contracts as "contracts that can be used to encode arbitrary state transition functions, as well as many others that we have not

yet imagined, simply by writing up the logic in a few lines of code." These contracts act as an agreement whose terms can be pre-programmed within a blockchain infrastructure with the ability to self-execute. The main goal of a smart contract is to enable two anonymous parties to trade and do business with each other, usually over the Internet, without the need for a middleman.

A smart contract works as follows:

- *Definition of the agreement*. A smart contract follows the same rules as a regular contract but executes the terms autonomously. For this reason, it is very important that the smart contract does exactly what parties agree to. Thus, it is necessary to define a precise set of conditions under which some actions take place;
- *Coding the contract*. The concrete definition of a contract is achieved by inputting the proper logic when writing the smart contract. The code behaves in predefined ways and does not have the linguistic nuances of human languages. It automatizes the "if this happens then do that" part of traditional contracts;
- *Uploading in the blockchain*. The code is encrypted and sent out to other computers via a distributed ledger network;
- *Execution of the contract*. The network updates the distributed ledger to record the execution of the contract, and then monitors for compliance within the terms of the smart contract. If the pre-arranged conditions happen, then the contract automatically executes.