Hyperledger Fabric basics

Structure and functionalities in depth

Clemente Serrano

November 2018

1 What is blockchain?

Blockchain is a point-to-point, decentralized and distributed network which operates publicly or privately under consensus models, digital signature, hash and commonly smart contracts. Within the network, the history of transactions that occur between members of the network is recorded, in the form of a sequential chain of packages or "blocks" linked together (hence the name blockchain) and unmodifiable (all information stored in the blockchain can not be edited or deleted).

The decentralization and distribution of the point-to-point network information prevents any member or group of members from controlling the entire system, they all keep an accurate and real-time copy of the transaction history. For its part, the consensus guarantees that the copies of the transaction history that each participant handles are accurate and with a low probability of being fraudulent (a fraud could occur if the records of all the members of the network are modified at the same time , which requires a high computing capacity, depending on the number of nodes). The hash functions applied in the opening of each block guarantee that any alteration in the information of some block generates a different hash value in all the blocks (each block hash is calculated with the hash and information of the previous block, therefore any modification of content in a block will be noticeably revealed in the rest). Digital signatures guarantee that transactions are made only by those who are within the network, not by external without authorization (each signature is asymmetrically encrypted). Finally, smart contracts allow a controlled and consistent access to the blockchain, defining the business rules by which transactions are governed.

In short, blockchain is a shared and replicated transaction system that is updated through smart contracts and is constantly synchronized through a collaborative process called consensus.

Basically, there are two types of blockchain: public and private. According to MONERIUM 2017 [1], within the classification of the blockchain are:

1. Public blockchains: A public blockchain does not have permission, anyone can send transactions and read the transaction book, and anyone can participate in the consensus. There is no central authority that maintains an official copy of the accounting book. The public blockchains are ”totally decentralized”.

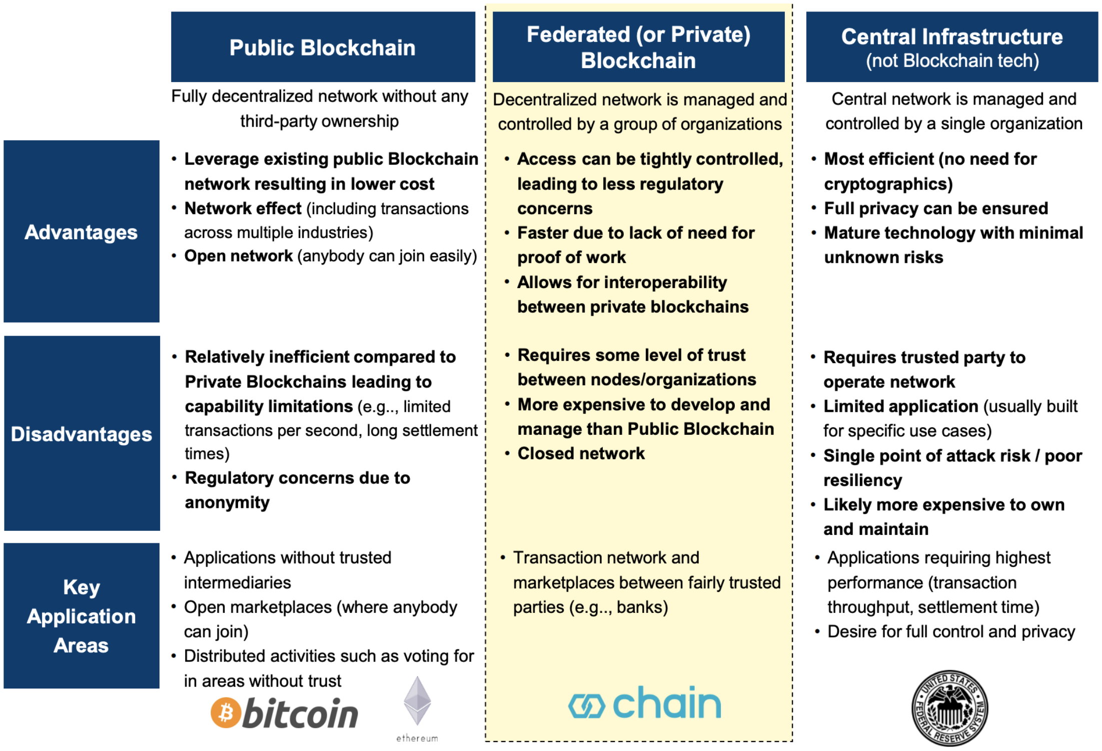
2. Consortium blockchains: a chain of consortium blocks is a chain of blocks where the consensus process is controlled by a set of pre-selected nodes; for example, a consortium of several financial institutions, each of which operates a node and most of which must sign each block to form a consensus. The right to read and write the chain of blocks may be restricted in part or in its entirety.

2 Hyperledger as a private blockchain

The Linux Foundation founded the Hyperledger project in 2015 to advance the blockchain technologies of the industry. Instead of declaring a single blockchain standard, it encourages a collaborative approach to develop blockchain technologies through a community process, with intellectual property rights that encourage open development and the adoption of key standards throughout the weather. [3]

Hyperledger Fabric is one of the blockchain projects within Hyperledger. Like other blockchain technologies, it has a ledger, uses smart contracts and is a system through which participants manage their transactions.

Where Hyperledger Fabric is separated from other blockchain systems is that it is private and authorized. Instead of an open system without permission that allows unknown identities to participate in the network (requiring protocols such as ”proof of work” to validate transactions and protect the network), members of a Hyperledger Fabric network are registered through a provider of reliable membership services.



	Public Blockchain	Federated (or Private) Blockchain	Central Infrastructure (not Blockchain tech)
	Fully decentralized network without any third-party ownership	Decentralized network is managed and controlled by a group of organizations	Central network is managed and controlled by a single organization
Advantages	• **Leverage existing public Blockchain network resulting in lower cost** • **Network effect** (including transactions across multiple industries) • **Open network** (anybody can join easily)	• **Access can be tightly controlled, leading to less regulatory concerns** • **Faster due to lack of need for proof of work** • **Allows for interoperability between private blockchains**	• **Most efficient (no need for cryptographics)** • **Full privacy can be ensured** • **Mature technology with minimal unknown risks**
Disadvantages	• **Relatively inefficient compared to Private Blockchains leading to capability limitations** (e.g.., limited transactions per second, long settlement times) • **Regulatory concerns due to anonymity**	• **Requires some level of trust between nodes/organizations** • **More expensive to develop and manage than Public Blockchain** • **Closed network**	• **Requires trusted party to operate network** • **Limited application** (usually built for specific use cases) • **Single point of attack risk / poor resiliency** • **Likely more expensive to own and maintain**
Key Application Areas	• Applications without trusted intermediaries • Open marketplaces (where anybody can join) • Distributed activities such as voting for in areas without trust	• Transaction network and marketplaces between fairly trusted parties (e.g.., banks)	• Applications requiring highest performance (transaction throughput, settlement time) • Desire for full control and privacy
	bitcoin ethereum	chain	UNITED STATES FEDERAL RESERVE SYSTEM

Figure 1: Types of blockchains, its advantages, disadvantages and areas of application [2].

Hyperledger Fabric also offers several pluggable options. The general ledger data can be stored in multiple formats, consensus mechanisms can be exchanged and released, and different membership service providers are allowed.

Hyperledger Fabric also offers the ability to create channels, which allows a group of participants to create a separate transaction book. This is an especially important option for networks in which some participants may be competitors and do not want each transaction they make to be known to all participants. If two participants form a channel, then those participants, and no other, have copies of that channel's accounting ledger.

3 Structure and operation of Hyperledger Fabric

3.1 Identity of participants

The Hyperledger blockchain network is made up of multiple actors, including approval nodes, ordering nodes, client applications, administrators and more. Each of these actors (active elements inside or outside the network that can consume their services) has a digital identity encapsulated in an X.509 digital certificate. These digital identities are extremely important since they determine the exact permissions on the resources and the access to the information that the actors have in a blockchain network.

For an identity to be verifiable, it must come from a trusted source. A membership service provider (MSP) is how this will be achieved in this project. More specifically, an MSP is a component that defines the rules that govern valid identities for an organization. The implementation of a default MSP uses X.509 certificates as identities, adopting a hierarchical public key infrastructure model (PKI) of its traditional Public Key Infrastructure. A PKI and an MSP work together: the PKI provides a list of identities and the MSP indicates which of them are members of a particular organization that participates in the network.

3.2 PKIs: Definition and components

A public-key infrastructure is a system of policies, procedures, entities, and services that support the use of asymmetric cryptography to provide secure communications over a network. It provides a framework to combine and achieve four main security functions in the network:

1. Authentication: Ability to verify the identity of a participant.
2. Confidentiality: Protection of information against unauthorized disclosure.
3. Integrity of data: Protection of information against unauthorized modifications.
4. Non-repudiation: Prevention of an entity denying actions already executed.

Although a blockchain network is more than a communications network, it uses the PKI standard to ensure secure communication between several network participants and to ensure that messages published in the block chain are authenticated correctly.

The main components of a PKI are the following:

1. **Digital certificates**: Digital document whose objective is to provide authentication to the participants of the network. It contains a set of attributes related to both the certificate holder and the issuer (who issues it). The most common type of certificate, and the one used in this project, is the one that complies with the X.509 standard. The common structure of an X.509 certificate contains multiple attributes, among which stand out: the general data of the certificate bearer, version of the certificate, serial number of the certificate, algorithm with which the identity of the bearer is encrypted, general data of the certifying entity that issued the certificate, lifetime of the certificate, public key of the certificate holder and digital signature of the issuer of the certificate (certifying authority).

   The attributes of each actor in the blockchain are registered in a digital certificate, which allows you to prove your identity in front of the system.

2. **Certification Authorities** (CAs): The digital identity with which an actor participates in the blockchain is issued by a trusted authority for the system. The certificates issued by a CA are digitally signed by it and link the participant of the network with its public key. As a result, if an actor trusts the CA (and knows his public key), he can trust that the specific stakeholder with whom he is interacting is linked to the public key included in the certificate and possesses the included attributes, by validating the signature of the CA in the certificate of the actor.

   As mentioned above, each actor that wants to interact with the blockchain needs an identity; it is the CA that provides the basis for the actors of an organization to have a verifiable digital identity.

   CAs come in two forms: root CA and intermediate CAs. Because root CAs must securely distribute hundreds of millions of certificates to users, it makes sense to spread this process through what is called intermediate CAs. These intermediate CAs have their certificates issued by the root CA or other intermediate authority, which allows the establishment of a ”chain of trust” for any certificate issued by any CA in the chain. This ability to track the root CA not only allows the role of CAs to escalate by providing security, allowing organizations that consume certificates to use intermediate CAs with confidence, it limits the exposure of the root CA, which, if compromised , could endanger the entire chain of trust. If an intermediate CA is compromised, on the other hand, there will be a much smaller exposure.

Intermediary CAs provide flexibility when it comes to issuing certificates in multiple organizations, which is useful in a private blockchain. For example, different organizations may use different root CAs or the same root CAs with different intermediate CAs; it depends on the needs of the network.

3. **Clients**: A PKI client is an entity that requests digital certificates from CAs and uses them to interact with the system. To obtain a digital certificate from a CA, a PKI client sends a public / private key generation request and then makes a request to issue a digital certificate (all of the above to the CA). Received the digital certificate, he uses it to identify himself as an actor of the system.

   The client is responsible for guaranteeing the security of his private key, since if he loses or reveals it, not only will he not be able to decipher the messages he receives, but he will also be able to be supplanted or intercepted by any other unauthorized actor.

4. **Certificate Revocation Lists** (CRL): A CRL is a public reading data structure signed by a CA which refers to those digital certificates that have been revoked in the system. Each certificate revoked by a CA is identified in a CRL by its serial number. When an actor makes use of a certificate (for example, to verify the digital signature of a remote user), the system not only verifies the validity and signature of the certificate, but also acquires a recently updated copy of the CRL and verifies that the The serial number of the certificate is not in it, that is, it is not revoked.

Digital certificates have a specific life time. Now, the circumstances that existed when the certificate was issued can change before the certificate expires naturally. Reasons for revocation include private key commitment, change of affiliation, name change, etc. (The specific reason codes are defined in X.509).

5. **Certificate Distribution System or Repository** (CDS): The CDS are used to provide mechanisms of storage of the certifications and the information of the CRLs. The term repository is usually associated with databases, however, in the context of a PKI, a repository is a generic term used to refer to any method to store and retrieve information related to PKI (mainly public key certificates and CRLs).

Now, how is it that these components interact with each other? What is specifically the mechanics of a PKI?

The operation of a PKI is based on five processes: issuance of certificates, revocation of certificates, creation and publication of certificate revocation lists, storage and retrieval of certificates and certificate revocation lists and finally management of the life cycle of the keys of encryption of information.

Before going into detail about these processes, it is necessary to understand how their operational core works: the digital signature.

3.3 PKIs: Digital signature for the integrity and authenticity of the information

The integrity of the information exchanged in a blockchain is a fundamental concept for its operation. When a participant of the network exchanges information with another participant, it is necessary to ensure that the information exchanged has not been modified during its transmission and that the sender of the message is who it claims to be.

The mechanisms of assurance of content integrity and authenticity in the network are based on digital signatures that, as the name suggests, allow a party to digitally sign their messages to guarantee the integrity of both the sender and the message sent.

A digital signature uses two cryptographic tools for its operation: hashing and asymmetric encryption. Next, each of them is defined roughly.

1. **Hash**: A hash is defined as a funtion $h : S \mapsto P$ with S : set of strings of random length and P : string of fixed length characters which must comply as much as possible with certain requirements that make it extremely useful for cryptography. These requirements are:

    (a) *Resistance to collisions*: It must be complex to find one or more texts that generate the same hash. In other words, it should be given as much as possible that $\forall s_a, s_b \in S : h(s_a) = h(s_b) \implies s_a = s_b$ (the hash function should be as close as possible to an injective function).

    (b) *One way* (first clause): It must be difficult to find $h^{-1}(p)$ (its inverse function). This means that, given a given image p_0, it must be complex to find a message s_0 such that $h(s_0) = p_0$.

    (c) *One way* (second clause): Given some pre image s_a, it must be complex to find a different pre image s_b that generates the same hash $h(s_a)$.

    (d) *Low cost*: Calculate $h(s)$ for any text s should involve a low temporal and spatial cost (processing time and machine memory).

    (e) *Uniformity*: Whatever the parameters of the function (keys), it must be equally probable to have a determined hash value, independently of any other element.

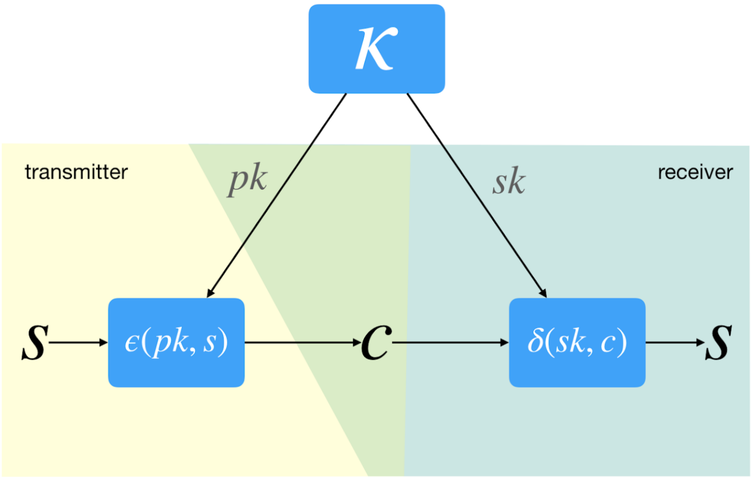
2. **Asymmetric encryption**: Asymmetric encryption is a method for the exchange of information between two parties by means of an encryption function that uses public and private parameters (called keys) of the participants to operate. The main characteristic of this method is that it is possible to encrypt and decipher by means of its parameters; one of the keys (the public one of the receiver) is used to encrypt, while the other (the private one of the same receiver) is used to decipher, which unlike hashing, makes the method a two-way procedure.

    An asymmetric encryption scheme $\Lambda(\kappa, \epsilon, \delta)$ consists of three algorithms:

    (a) *Algorithm for the generation of random keys* κ, which does not receive inputs and returns a public and private key coordinate (pk, sk). Both the private key and the public key correspond to the recipient of the message, however the private only he knows it and public it, the issuer and any other actor that is in the system.

(b) *Encryption algorithm* ϵ, which takes the public key pk of the receiver and a string of characters s (message to be transmitted, commonly called plain text) to return the encrypted message c (commonly called encrypted text).

(c) *Decryption algorithm* δ, which takes the private key sk from the receiver and the encrypted message c, to return the original message s.



© 2018 Clemente Serrano. All rights reserved.

Figure 2: Scheme of operation of the asymmetric encryption. Source: self made.

With clarity regarding the concepts of hashing and asymmetric encryption, one can go on to explain what a digital signature is and what its role is in an authentication process within a PKI.

A digital signature is a digital code generated and authenticated by hashing and asymmetric encryption which is attached to an electronically transmitted document. Before sending a message, the sender generates a hash with its content, which it then sequentially encrypts twice: a first using its private key and a second one with the public key of the receiver. The generated code is the digital signature, which is attached to the message and then sent. When the recipient receives the message, he uses his private key to decrypt the issuer's signature, and then the issuer's public key in order to obtain the content's hash. Finally, using the same hash function as the issuer, it encodes the content of the message and compares it with the attachment in the issuer's signature.

If the generated hash is the same as that of the signature, then the integrity of the message is verified, otherwise the message has been adulterated (one of the properties of the hash function $h(s) = p$ is that a minimum change in s generates a noticeable change in p, which allows to easily identify adulterations in the content of messages). Figure 3 shows the operation of a digital signature graphically.

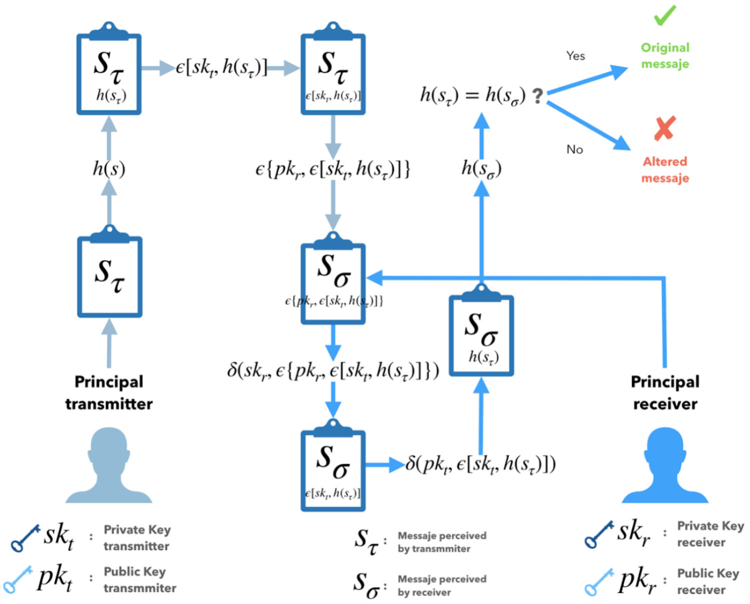

S_τ
$h(s_\tau)$
$\epsilon[sk_t, h(s_\tau)]$
S_τ
$\epsilon[sk_t, h(s_\tau)]$
$h(s_\tau) = h(s_\sigma)$?
Yes
Original messaje
No
Altered messaje
$h(s)$
$\epsilon\{pk_r, \epsilon[sk_t, h(s_\tau)]\}$
$h(s_\sigma)$
S_τ
S_σ
$\epsilon\{pk_r, \epsilon[sk_t, h(s_\tau)]\}$
S_σ
$h(s_\tau)$
$\delta(sk_r, \epsilon\{pk_r, \epsilon[sk_t, h(s_\tau)]\})$
Principal transmitter
Principal receiver
S_σ
$\epsilon[sk_t, h(s_\tau)]$
$\delta(pk_t, \epsilon[sk_t, h(s_\tau)])$
sk_t : Private Key transmitter
pk_t : Public Key transmmiter
S_τ : Messaje perceived by transmmiter
S_σ : Messaje perceived by receiver
sk_r : Private Key receiver
pk_r : Public Key receiver

Figure 3: Scheme of operation of the digital signature. Source: self made.

3.4 PKIs: Operation