Blockchain explained

by Jerome Kehrli

Witten on Friday Oct 07, 2016

The blockchain and blockchain related topics are becoming increasingly discussed and studied nowadays. There is not one single day where I don't hear about it, that being on linkedin or elsewhere.
I interested myself deeply in the blockchain topic recently and this is the first article of a coming whole serie around the blockchain.

This article presents an introduction on the blockchain, presents what it is in the light of its initial deployment in the Bitcoin project as well as all technical details and architecture concerns behind it.
We won't focus here on business applications aside from what is required to present the blockchain purpose, more concrete business applications and evolutions will be the topic of another post in the coming days / weeks.

This article presents and explains all the key techniques and mechanisms behind the blockchain technology and is divided in the following sections :.

One might want to see part of this article as a slideshare presentation available here :http://www.slideshare.net/JrmeKehrli/the-blockchain-the-technology-behind-bitcoin.

Table of Contents

The blockchain principles and fundamentals are really coming initially from the design work on the Bitcoin. Most of this article focuses on the design and the principle of the blockchain put in place in the Bitcoin system.
Some more recent (Blockchain 2.0) implementations differ slightly while still sharing most genes with the original blockchain, making all that is presented below valid from a conceptual perspective in these other implementations as well.

1. What is the blockchain ?

1.1 Some definitions

Answering this question is somewhat tricky. I will give three definitions that do a pretty good job in answering to this question.

Initial Definition :

The blockchain is the technology running the bitcoin.

The blockchain is a technology that underlies bitcoin - conceived in 2008 and first implemented in 2009 - where it serves as the public ledger for all transactions. The blockchain technology has been conceived bySatoshi Nakamoto a virtual identity which is believed nowadays to belong to a group rather than a single individual.

Wikipedia's definition :

A blockchain is a distributed database that maintains a continuously growing list of records called blocks secured from tampering and revision.

A blockchain consists of blocks that hold batches of valid transactions. Each block includes the hash of the prior block in the blockchain, linking the two. The linked blocks form a chain.
In addition to a secure hash based history, any blockchain database has a specified algorithm for scoring different versions of the history so that one with a higher value

can be selected over others. Peers supporting the database don't have the exact same version of the history at all times, but fall to eventual consistency.

My definition, slightly different from wikipedia's, which underlines what are, from my perspective, the key aspects of the blockchain *technology*, as opposed to the blockchain *data structure* on which wikipedia's definition focuses:

The blockchain is a secured protocol enabling peer-to-peer exchanges on a distributed network in a secured, public and non-repudiable way.

I guess these three definitions are very valid and the three of them give some important information. I do however prefer mine (well that's a surprise ...) since it underlines what I believe are the very key aspects of the blockchain.
The blockchain is first and foremost a certification infrastructure which would benefit all applications relying on it.
First the blockchain is more an application protocol than anything else since it consists if individual behaviour specification that eventually lead to a distributed, peer-to-peer, message broadcast-based and secured information database.

1.2 A tiny little bit of history

The blockchain architecture and principles was first designed for bitcoin, as a solution to the problem of making a database both secured and widely distributed. The block chain is the main innovation of Bitcoin.

As of 2014, "*Blockchain 2.0*" was a term used in the distributed blockchain database field. Blochchain 2.0 is an evolution of the initial blockchain intent where pretty much only transaction from a sender to a receiver could be stored. In this new paradigm, instead of simple transactions, the exchanges happen around so called "*Smart Contracts*", actually pretty complete applications implemented in specific scripting languages.

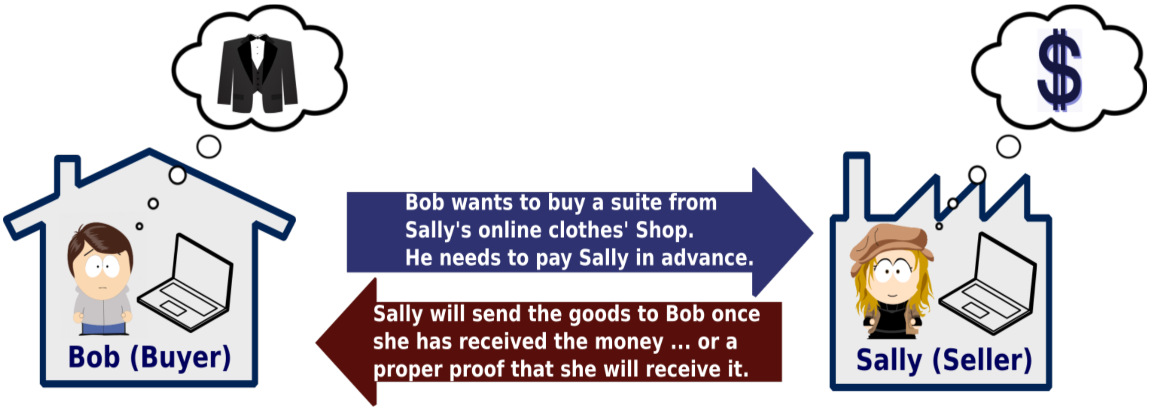
The Economist described one implementation of this second-generation programmable blockchain as coming with "*a programming language that allows users to write more sophisticated smart contracts, thus creating invoices that pay themselves when a shipment arrives or share certificates which automatically send their owners dividends if profits reach a certain level.*"

1.3 Introduction example

Let's imagine the following example which we will be using throughout this article to illustrate the blockchain principles :

- Bob is an online buyer who just discovered Sally's online clothes shop. He found a nice suite there and wants to buy it.

- Sally is running a little online shop and sells various kinds of clothes. Sally's little shop only accepts paypal transactions.
- Bob needs to pay Sally, using paypal, before she sends him the suite .


Bob (Buyer)
Bob wants to buy a suite from Sally's online clothes' Shop. He needs to pay Sally in advance.
Sally will send the goods to Bob once she has received the money ... or a proper proof that she will receive it.
Sally (Seller)

2. Key problem solved by the blockchain

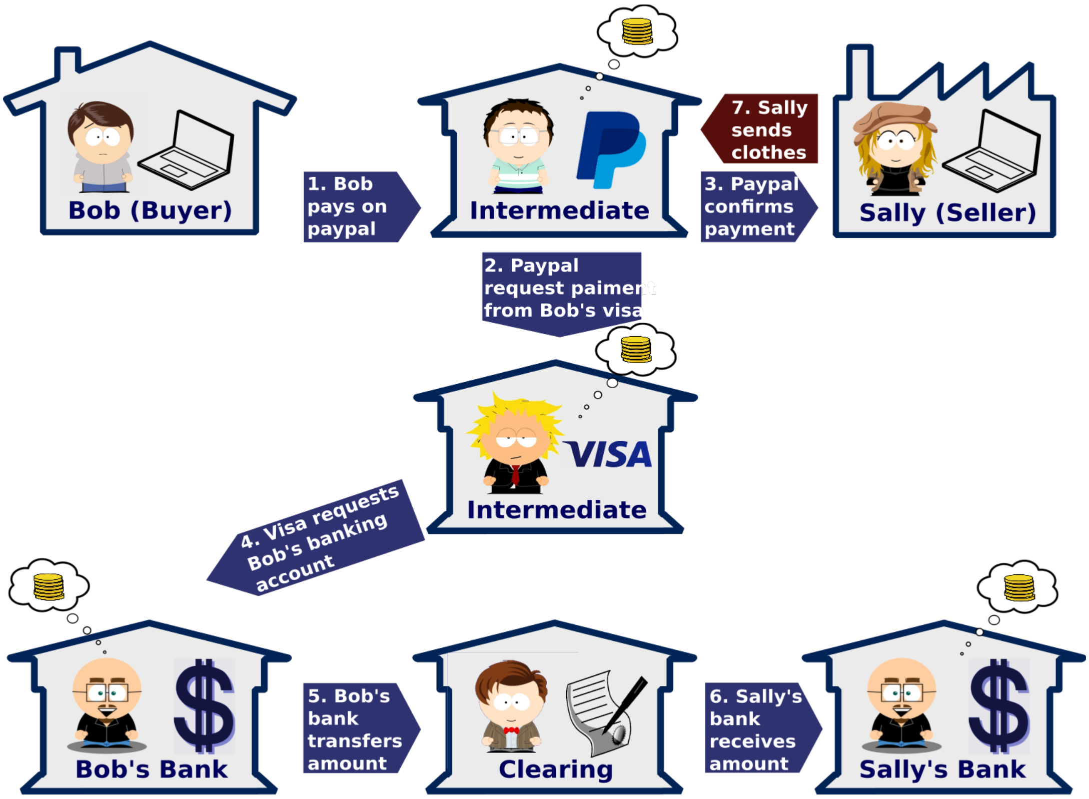
Banking is a system of intermediaries across the spectrum - ranging from payment networks (e.g Mastercard, Visa etc) to Clearinghouses in Capital Markets to Banks, etc. And the reason these intermediaries exist is to establish trust between two parties who do not know each other.

The Blockchain stands behind every bitcoin ever created by maintaining the proof of ownership.
The innovation is it's openness yet it's security which ensures that the currency is tamper proof.

2.1 Back on the Introduction example.

Recall Bob and Sally's transaction above ?

This is how it happens in practice, with all the intermediates


Bob (Buyer)
1. Bob pays on paypal
Intermediate
7. Sally sends clothes
3. Paypal confirms payment
Sally (Seller)
2. Paypal request paiment from Bob's visa
VISA
Intermediate
4. Visa requests Bob's banking account
Bob's Bank
5. Bob's bank transfers amount
Clearing
6. Sally's bank receives amount
Sally's Bank

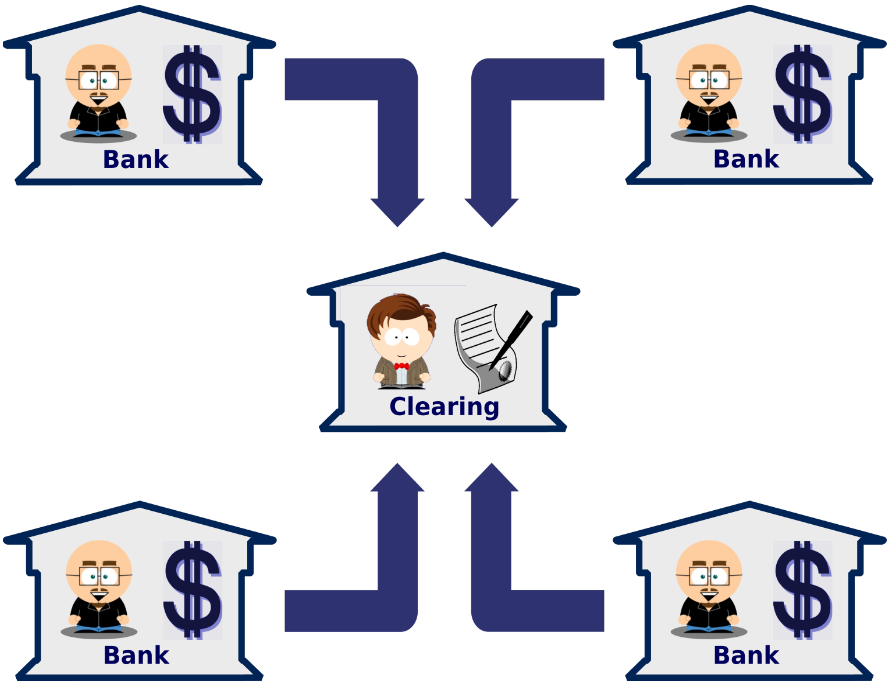
This usual model suffers from several problems / drawbacks:

- The financial system is opaque and lacks transparency and fairness.
- All these intermediates are no volunteers. They work for money and get paid for their services. The transaction costs money to both the buyer and the seller. There are interest rates, fees, surcharges, etc. EFTs in Europe can cost 25 euros. Credit transactions can cost several percent of the transaction.
- All these exchanges are error prone. Credit card informations are stolen. Banks make mistakes.
- An account holder is eventually not even the actual owner of his account. The bank really owns the account. Funds can be garnished, even frozen completely.
- Banks and other payment processors like PayPal, Visa, and Mastercard may refuse to process payments for certain legal entities.
- Financial exchanges are slow. Checking and low cost wire services take days to complete.

2.2 Centralization and Clearing Houses

A clearing house is a financial institution that provides clearing and settlement services for financial and commodities derivatives and securities transactions. These transactions may be executed on a futures exchange or securities exchange, as well as off-exchange in the over-the-counter (OTC) market.
A clearing house stands between two clearing firms (also known as member firms or clearing participants) and its purpose is to reduce the risk of one (or more) clearing firm failing to honor its trade settlement obligations. A clearing house reduces the settlement risks by netting offsetting transactions between multiple counterparties.

Buyers and sellers use intermediaries because they may not trust the other party, but they **trust that the intermediary will assure the transaction is completed faithfully**. This is the fundamental role of a clearing house as illustrated below:



The clearing house provides protection to the problems:

- How do you ensure some fund will not be spent twice ?
- How do you ensure some the transaction sender actually has the funds ?

In addition, the clearing house holds the central **transaction ledger**.

The problem with these central ledgers, or clearing houses

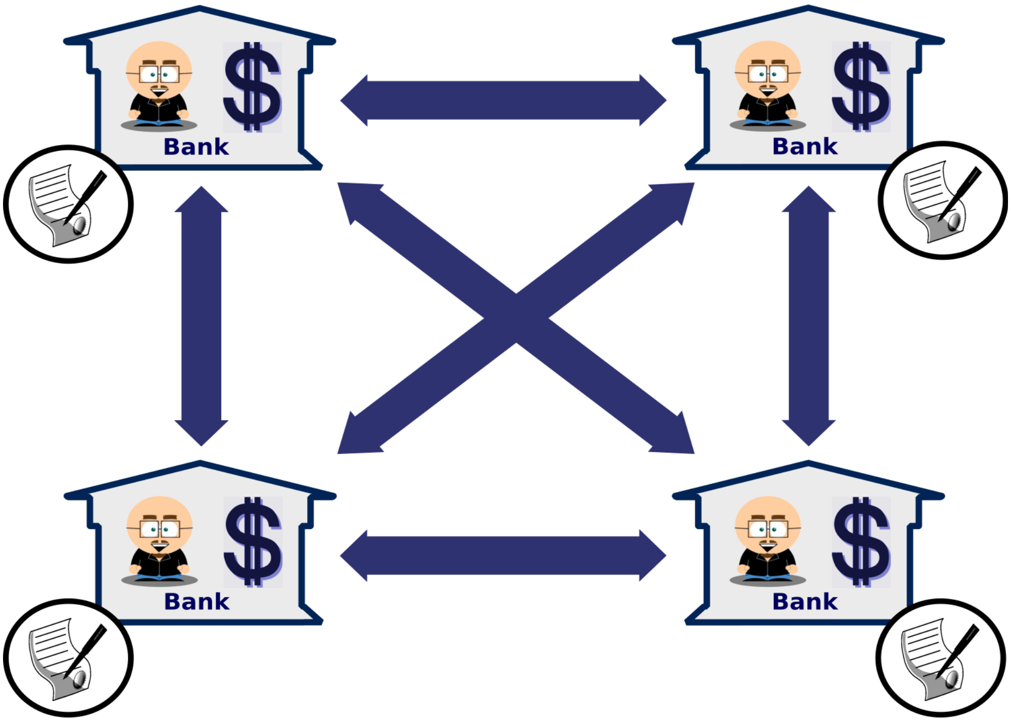
When one bank sends money to another, no physical currency changes hands. Banks and settlement systems use central electronic ledgers to track assets. But such central ledgers - or clearing houses - can be slow and inefficient, often relying on faxes or manual input.

That not only wastes time but racks up fees. The system is also open to hacking and fraud. These central institutions gets fees to cover such risks of course as well as many other services, bue the price is high and prevents, for instance, micro-paiements who are not able to support the charge asked by these central structures.

2.3 Here comes the blockchain

In contrast to today's networks, **distributed ledgers** eliminate the need for central authorities to certify ownership and clear transactions. They can be open, verifying anonymous actors in the network, or they can be closed and require actors in the network to be already identified.
The best known existing use for the distributed ledgers is the cryptocurrency bitcoin.

Eliminating the need of the central ledger is precisely the core intent of the blockchain and all protocols put in place there adresse the need of making it public, permanent, distributed and secure.



A **distributed ledger** (also called **shared ledger**) is a consensus of replicated, shared, and synchronized digital data geographically spread across multiple sites, countries, and/or institutions.
Every node in the decentralized system has a copy of the blockchain. No centralized "official" copy exists and no user is "trusted" more than any other.

Users of Distributed Ledger Technology (DLT) significantly benefit from the efficiencies and economics by creating a more robust environment for real-time and secure data sharing. Contrary to common belief, the Bitcoin blockchain is not the

only distributed ledger, in fact, many other users of Distributed Ledger Technology use different methodologies to achieve the same consensus (e.g. Ripple, MultiChain, HyperLedger Project).
A blockchain is mostly a distributed ledger but not all distributed ledgers are blockchains. Although the term "blockchain" is used more frequently than "distributed ledger" in discussions, a blockchain is only one of the many types of data structures that provide secure and valid achievement of distributed consensus.
The bitcoin blockchain, and similar bockchains, which uses "Proof-of-Work" mining, is the most publicly proven method used to achieve distributed consensus

This leads us to another definition of the blockchain :

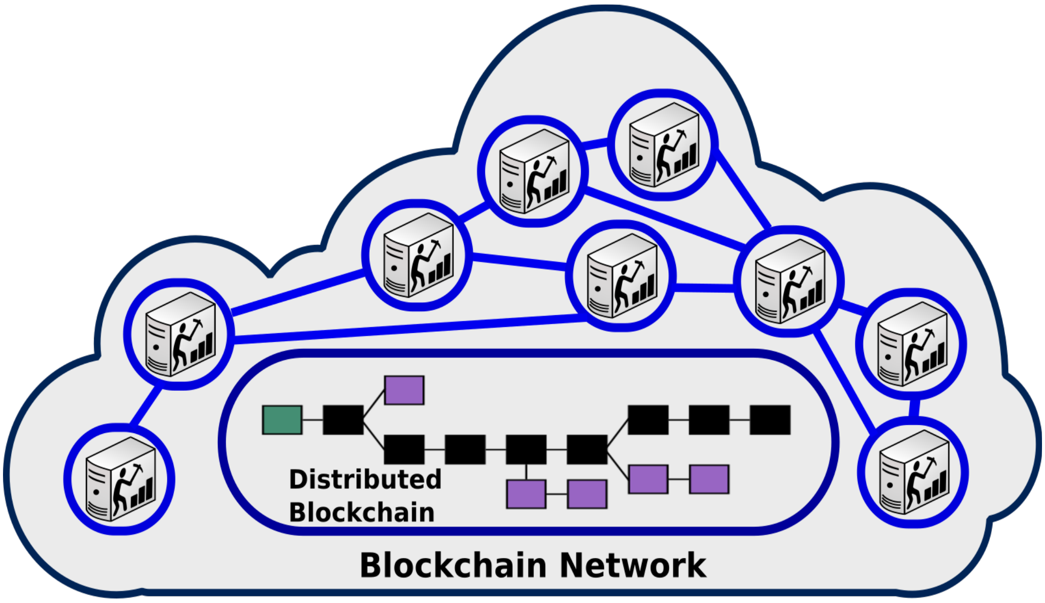
A blockchain is a type of distributed ledger, comprised of unchangeable, digitally recorded data in packages called blocks.

3. Operation of the blockchain

3.1 A simplified view of the blockchain

When discussing the Blockchain technology, one is really mentioning the blockchain itself and the network built around it as well as all protocols involved.

All of this can be represented this way, under a simplified form:



Key aspects here are :

- The blockchain network is a peer-to-peer network of independent nodes communicating together by message broadcasting.

- The key component of the network is the blockchain. **Every node has its own copy of the blockchain**
- A node is not necessarily connected to every other node, but at least some of them.

The blockchain itself is a list of blocks. These digitally recorded "blocks" of data are stored in a linear chain. Each block in the chain contains data (e.g. bitcoin transaction) and is cryptographically hashed.
Each block includes the hash of the prior block in the blockchain, linking the two, ensuring all data in the overall "blockchain" has not been tampered with and remains unchanged.

This has the effect of creating a chain of blocks from the genesis block to the current block. Each block is guaranteed to come after the previous block chronologically because the previous block's hash would otherwise not be known.
Each block is also computationally impractical to modify once it has been in the chain for a while because every block after it would also have to be regenerated. **The linked blocks form a chain**.

3.2 The bitcoin blockchain

A block chain is a transaction database shared by all nodes participating in a system based on the Bitcoin protocol. A full copy of a currency's block chain contains every transaction ever executed in the currency. With this information, one can find out how much value belonged to each address at any point in history.

For any block on the chain, there is only one path to the genesis block. Coming from the genesis block, however, there can be forks.
One-block forks are created from time to time when two blocks are created just a few seconds apart. When that happens, generating nodes build onto whichever one of the blocks they received first. Whichever block ends up being included in the next block becomes part of the main chain because that chain is longer.

The block chain is broadcasted to all nodes on the networking using a flood protocol.

3.3 Operation Principle Overview

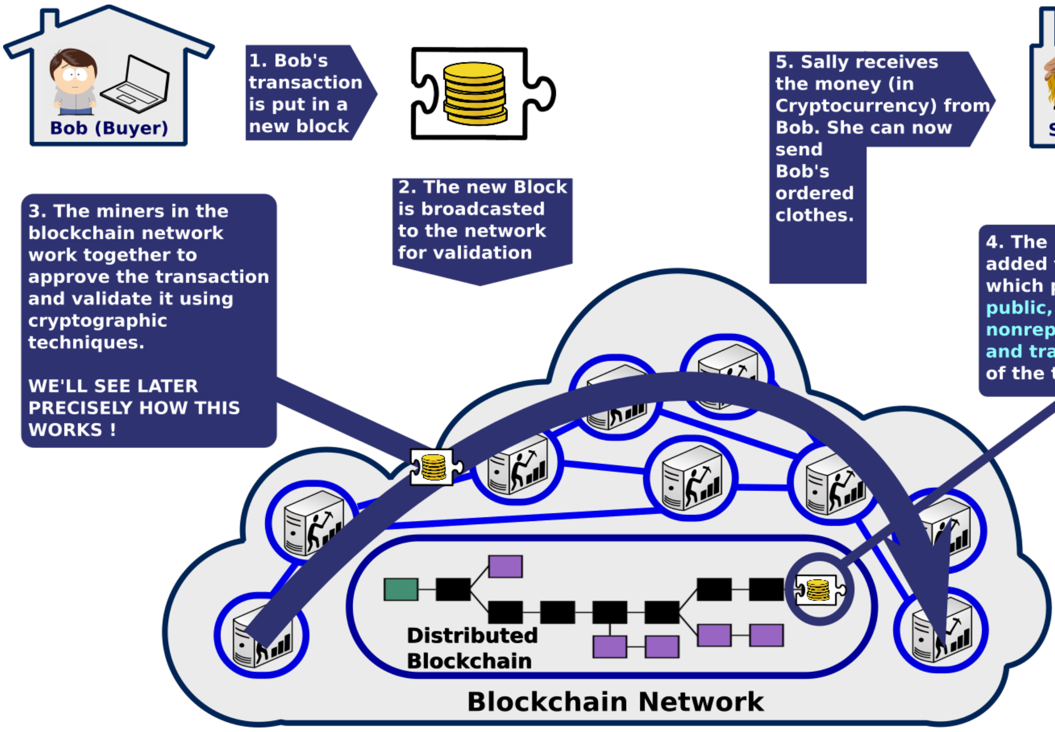
The intial blockchain, running the bitcoin system, provides distributed, public and secured storage of bitcoin transactions.

The operation principle of is pretty straightforward to understand:

- A user wants to pay another user some bitcoins, he broadcasts a transaction to the network.
- Miners add the transaction as they receive it to their current block, the one they are currently working on

- Randomly, one of the miner may win the lottery and "*mine*" the block (we'll get back to that)
- At that moment, this new "*definitive*" block is broadcasted to the network and added to everyone's copy of the blockchain

This is illustrated in the following visualization:


Bob (Buyer)
1. Bob's transaction is put in a new block
2. The new Block is broadcasted to the network for validation
3. The miners in the blockchain network work together to approve the transaction and validate it using cryptographic techniques.
WE'LL SEE LATER PRECISELY HOW THIS WORKS !
4. The Block is then added to the Chain which provides public, permanent, nonrepudiable, and transparent record of the transaction !
5. Sally receives the money (in Cryptocurrency) from Bob. She can now send Bob's ordered clothes.
Sally (Seller)
Distributed Blockchain
Blockchain Network

3.4 Miners and the "Proof of Work"

In order for a block to be accepted by network participants, miners must complete a *proof of work* which covers all of the data in the block.
The difficulty of this work is adjusted so as to limit the rate at which new blocks can be generated by the network to one every 10 minutes. Due to the very low probability of successful generation, this makes it unpredictable which worker computer in the network will be able to generate the next block.

The *proof of work* is a piece of data which is difficult (costly, time-consuming) to produce but easy for others to verify and which satisfies certain requirements. Producing a proof of work can be a random process with low probability so that a lot of trial and error is required on average before a valid proof of work is generated.

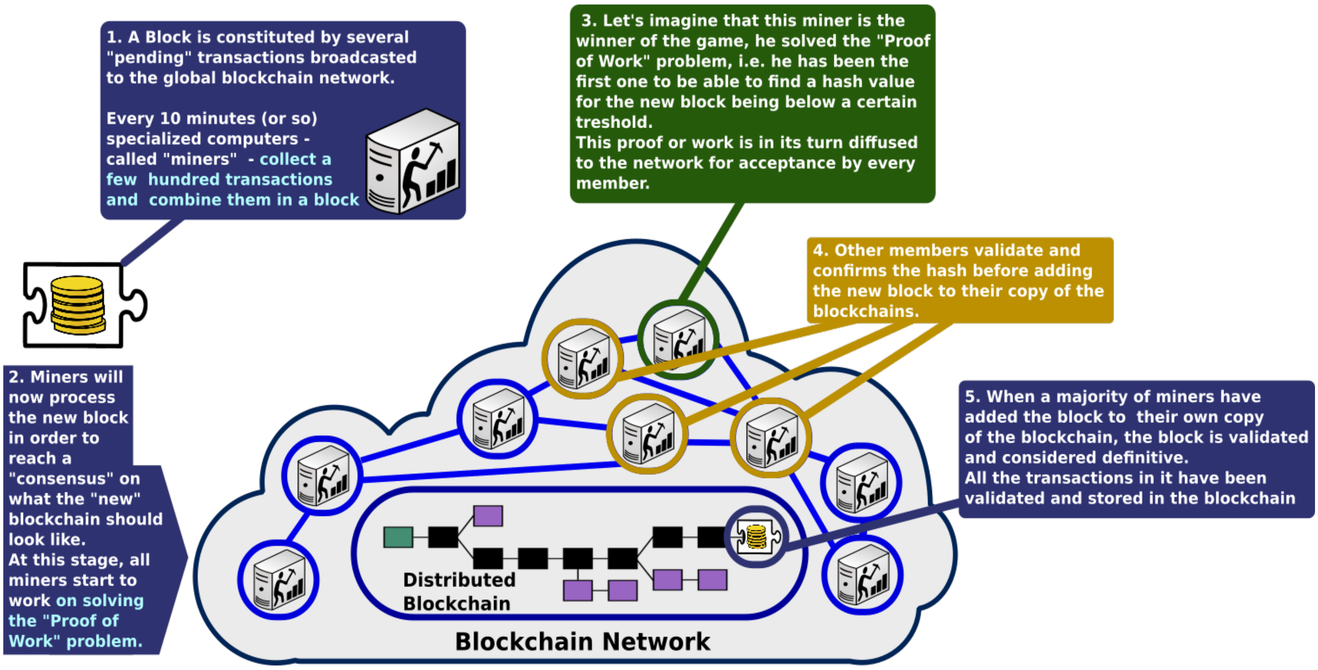
Bitcoin uses the Hashcash proof of work system.

For a block to be valid it must hash to a value less than the current target; this means that each block indicates that work has been done generating it. Each block contains the hash of the preceding block, thus each block has a chain of blocks that

together contain a large amount of work.
Changing a block (which can only be done by making a new block containing the same predecessor) requires regenerating all successors and redoing the work they contain. This protects the block chain from tampering.

The most widely used proof-of-work scheme is based on SHA-256 and was introduced as a part of Bitcoin.

The mining process works this way:



1. A Block is constituted by several "pending" transactions broadcasted to the global blockchain network.

Every 10 minutes (or so) specialized computers - called "miners" - collect a few hundred transactions and combine them in a block

2. Miners will now process the new block in order to reach a "consensus" on what the "new" blockchain should look like.
At this stage, all miners start to work on solving the "Proof of Work" problem.

3. Let's imagine that this miner is the winner of the game, he solved the "Proof of Work" problem, i.e. he has been the first one to be able to find a hash value for the new block being below a certain treshold.
This proof or work is in its turn diffused to the network for acceptance by every member.

4. Other members validate and confirms the hash before adding the new block to their copy of the blockchains.

5. When a majority of miners have added the block to their own copy of the blockchain, the block is validated and considered definitive.
All the transactions in it have been validated and stored in the blockchain

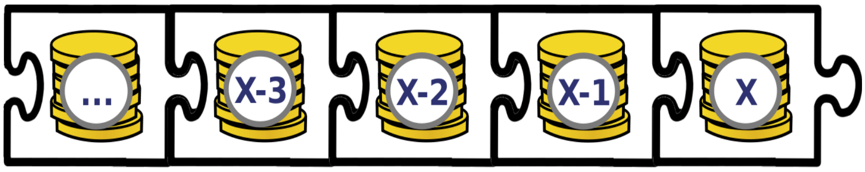
4. Technical aspects of the blockchain

Initially in the bitcoin system, a block chain is a **transaction database** shared by all nodes participating in a system based on the Bitcoin protocol.
A full copy of a currency's (bitcoin) lock chain contains every transaction ever executed in the currency. With this information, one can find out how much value belonged to each address at any point in history.

4.1 The Blockchain Data Structure

The blockchain data structure is an ordered, back-linked list of blocks of transactions. Every block contains a hash of the previous block. This has the effect of creating a chain of blocks from the genesis block to the current block.
Each block is guaranteed to come after the previous block chronologically because the previous block's hash would otherwise not be known.

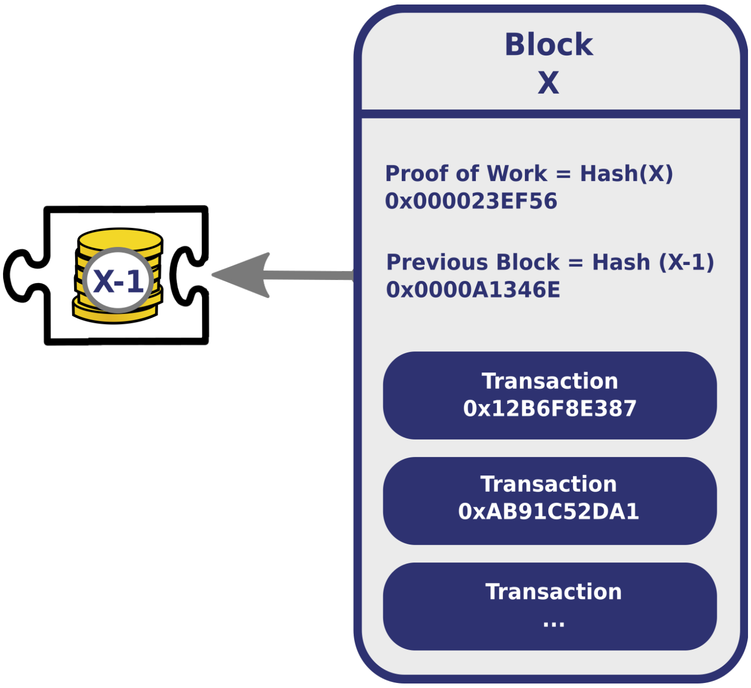
Each block is also computationally impractical to modify once it has been in the chain for a while because every block after it would also have to be regenerated.



Transaction data is permanently recorded in files called blocks.
They can be thought of as the individual pages of a city recorder's recordbook (where changes to title to real estate are recorded) or a stock transaction ledger.
Blocks are organized into a linear sequence over time (also known as the block chain).
New transactions are constantly being processes by miners into new blocks which are added to the end of the chain and can never be changed or removed once accepted by the network.

4.2 A first view on a block structure

Each block contains, among other things, a record of some or all recent transactions, and a reference to the block that came immediately before it. It also contains an answer to a difficult-to-solve mathematical puzzle , the hash or "Proof of Work".



4.3 A miner's life

In the Bitcoin world, transactions are broadcast to the network by the sender, and all peers trying to solve blocks collect the transaction records and add them to the block they are working to solve. This is called**Mining**.

Mining is the process of adding transaction records to Bitcoin's public ledger of past transactions. This ledger of past transactions is called the block chain as it is a chain of blocks.
The block chain serves to confirm transactions to the rest of the network as having taken place. Bitcoin nodes use the block chain to distinguish legitimate Bitcoin transactions from attempts to re-spend coins that have already been spent elsewhere.

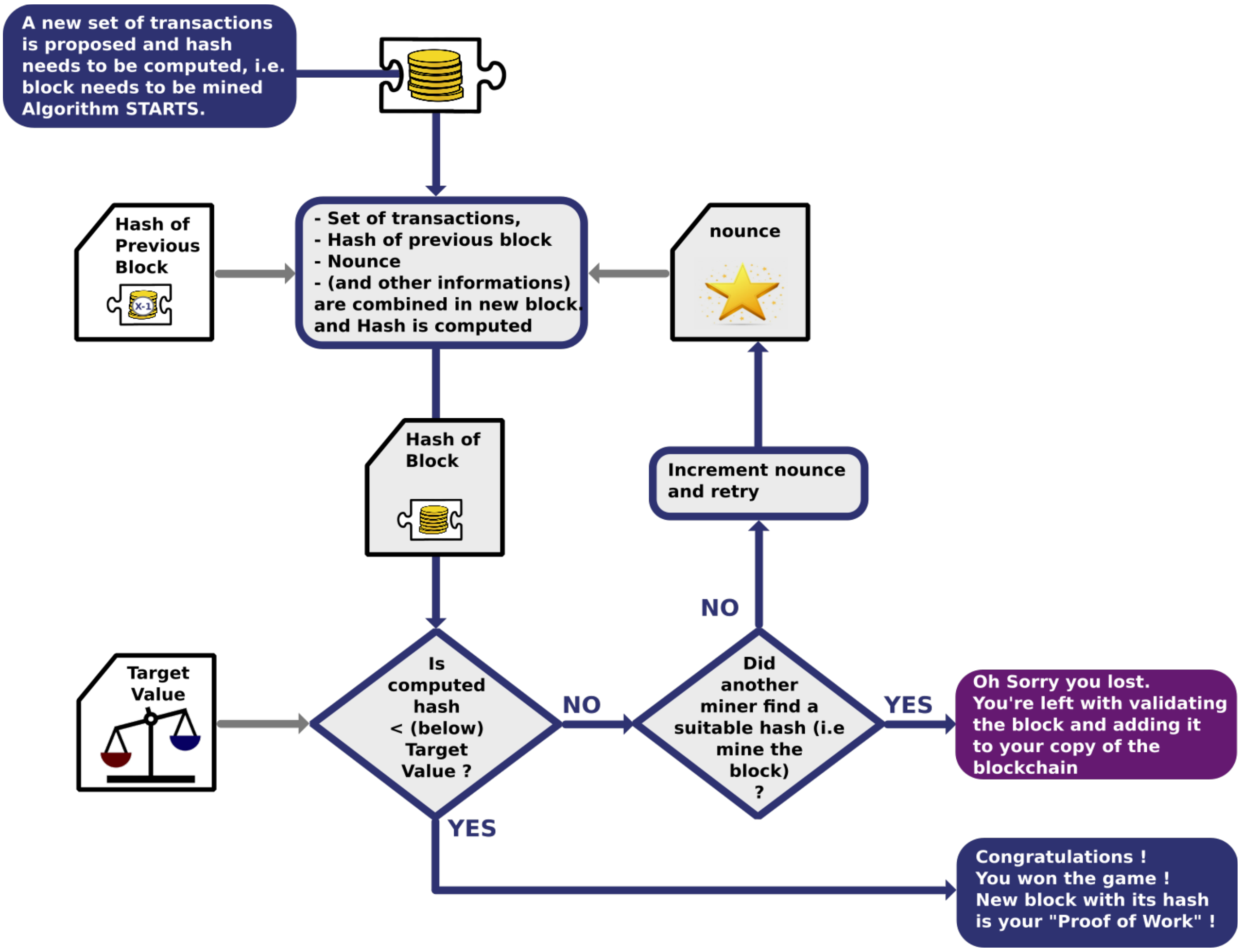
Mining is intentionally designed to be resource-intensive and difficult so that the number of blocks found each day by miners remains steady. Individual blocks must contain a **proof of work** to be considered valid.
This proof of work is verified by other Bitcoin nodes each time they receive a block. Bitcoin uses the hashcash proof-of-work function.

The primary purpose of mining is to allow Bitcoin nodes to reach a secure, tamper-resistant **consensus**.

The algorithm

Mining a block is difficult because the SHA-256 hash of a block's header must be lower than or equal to the target in order for the block to be accepted by the network.
This problem can be simplified for explanation purposes: The hash of a block must start with a certain number of zeros. The probability of calculating a hash that starts with many zeros is very low, therefore many attempts must be made.
In order to generate a new hash each round, a nonce is incremented.

Miners implement following (simplified) algorithm :


A new set of transactions is proposed and hash needs to be computed, i.e. block needs to be mined Algorithm STARTS.
Hash of Previous Block
X-1
- Set of transactions,
- Hash of previous block
- Nounce
- (and other informations)
are combined in new block. and Hash is computed
nounce
Hash of Block
Increment nounce and retry
NO
Target Value
Is computed hash < (below) Target Value ?
NO
Did another miner find a suitable hash (i.e mine the block) ?
YES
Oh Sorry you lost. You're left with validating the block and adding it to your copy of the blockchain
YES
Congratulations ! You won the game ! New block with its hash is your "Proof of Work" !

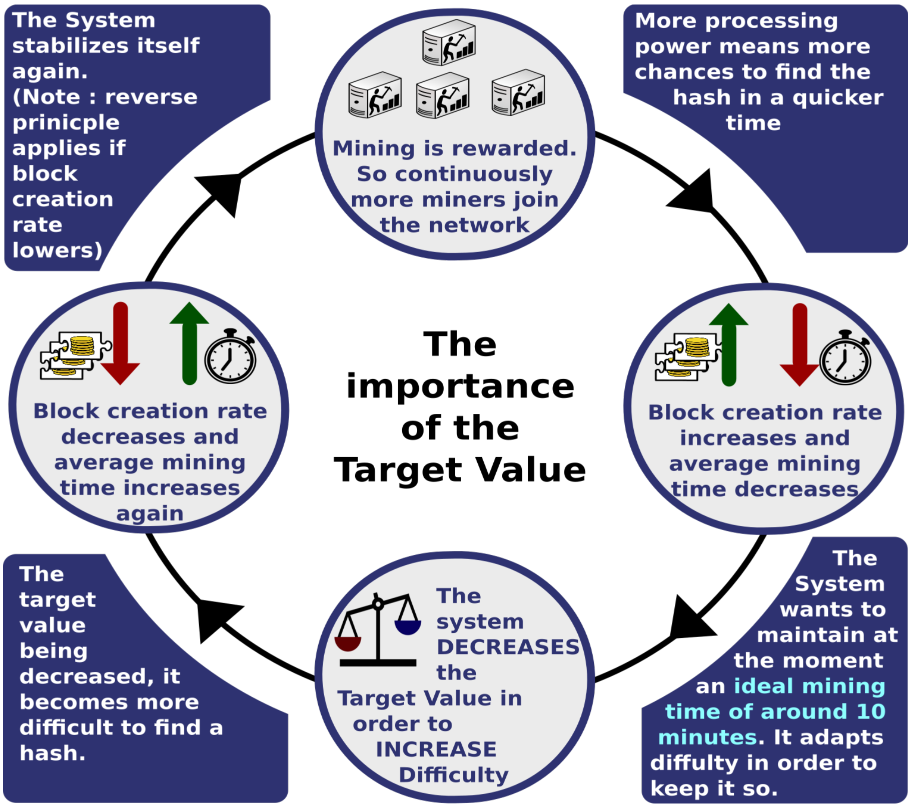
Note 2016.11.21: I should enhance the above schema to illustrate the fact that when a round trip fails, the hash could not be computed, after incrementing the nounce but before retrying, miners add to the current block they are working on the new transactions that they might have received in the meantime. Thus every round occurs potentially on a slightly different block. This is important because it gives its rationality to using Merkle Trees to compute the hash.

4.4 Difficulty Adjustment

The difficulty is the measure of how difficult it is to find a new block compared to the easiest it can ever be. It is recalculated every 2016 blocks to a value such that the previous 2016 blocks would have been generated in exactly two weeks had everyone been mining at this difficulty. This will yield, on average, one block every ten minutes.

As more miners join, the rate of block creation will go up. As the rate of block generation goes up, the difficulty rises to compensate which will push the rate of block creation back down.

Any blocks released by malicious miners that do not meet the required difficulty target will simply be rejected by everyone on the network and thus will be worthless.


The importance of the Target Value
Mining is rewarded. So continuously more miners join the network
More processing power means more chances to find the hash in a quicker time
Block creation rate increases and average mining time decreases
The System wants to maintain at the moment an ideal mining time of around 10 minutes. It adapts diffulty in order to keep it so.
The system DECREASES the Target Value in order to INCREASE Difficulty
The target value being decreased, it becomes more difficult to find a hash.
Block creation rate decreases and average mining time increases again
The System stabilizes itself again. (Note : reverse principle applies if block creation rate lowers)

Again, the difficulty of the mathematical problem is automatically adjusted by the network, such that it targets a goal of solving an average of 6 blocks per hour. Every 2016 blocks (solved in about two weeks), all Bitcoin clients compare the actual number created with this goal and modify the target by the percentage that it varied.

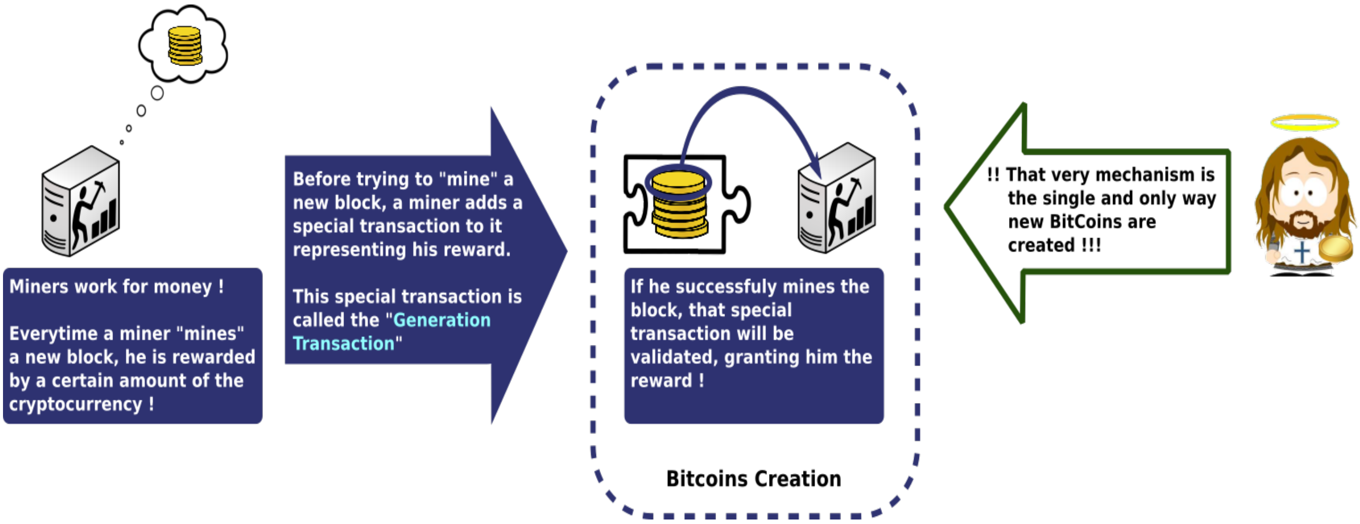
The network comes to a consensus and automatically increases (or decreases) the difficulty of generating blocks.

4.5 Miner retribution (and bitcoin creation)

Mining is also the mechanism used to introduce Bitcoins into the system: Miners are paid any transaction fees as well as a "subsidy" of newly created coins. This both serves the purpose of disseminating new coins in a decentralized manner as well as motivating people to provide security for the system.
It gives miners incentive to put their computation power at the disposal of the blockchain network.

Because there is a reward of brand new bitcoins for solving each block, every block also contains a record of which Bitcoin addresses or scripts are entitled to receive the reward. This record is known as a **generation transaction** (or a coinbase transaction) and is always the first transaction appearing in every block.

The number of Bitcoins generated per block starts at 50 and is halved every 210,000 blocks (about four years).



Miners work for money !

Everytime a miner "mines" a new block, he is rewarded by a certain amount of the cryptocurrency !

Before trying to "mine" a new block, a miner adds a special transaction to it representing his reward.

This special transaction is called the "Generation Transaction"

If he successfuly mines the block, that special transaction will be validated, granting him the reward !

Bitcoins Creation

!! That very mechanism is the single and only way new BitCoins are created !!!

In addition to the generation transaction, Miners get incentive to include transactions in their blocks because of attached transaction fees. A fee (pretty little) is perceived form every transaction in the newly mined block.

4.6 Bitcoin limited supply

In the specific case of the bitcoin, Satoshi had very soon the idea of limiting the bitcoin supply. There is an important reason of course behind this.

In a centralized economy, currency is issued by a central bank at a rate that is supposed to match the growth of the amount of goods that are exchanged so that these goods can be traded with stable prices. The monetary base is controlled by a central bank. In the United States, the Fed increases the monetary base by issuing currency, increasing the amount banks have on reserve, and more recently, printing money electronically in a process called Quantitative Easing.

In a fully decentralized monetary system, there is no central authority that regulates the monetary base. Instead, currency is created by the nodes of a peer-to-peer network. The Bitcoin generation algorithm defines, in advance, how currency will be created and at what rate. Any currency that is generated by a malicious user that does not follow the rules will be rejected by the network and thus is worthless.

Bitcoins are created each time a user discovers a new block. The rate of block creation is adjusted every 2016 blocks to aim for a constant two week adjustment period (equivalent to 6 per hour.) The number of bitcoins generated per block is set to decrease geometrically, with a 50% reduction every 210,000 blocks, or

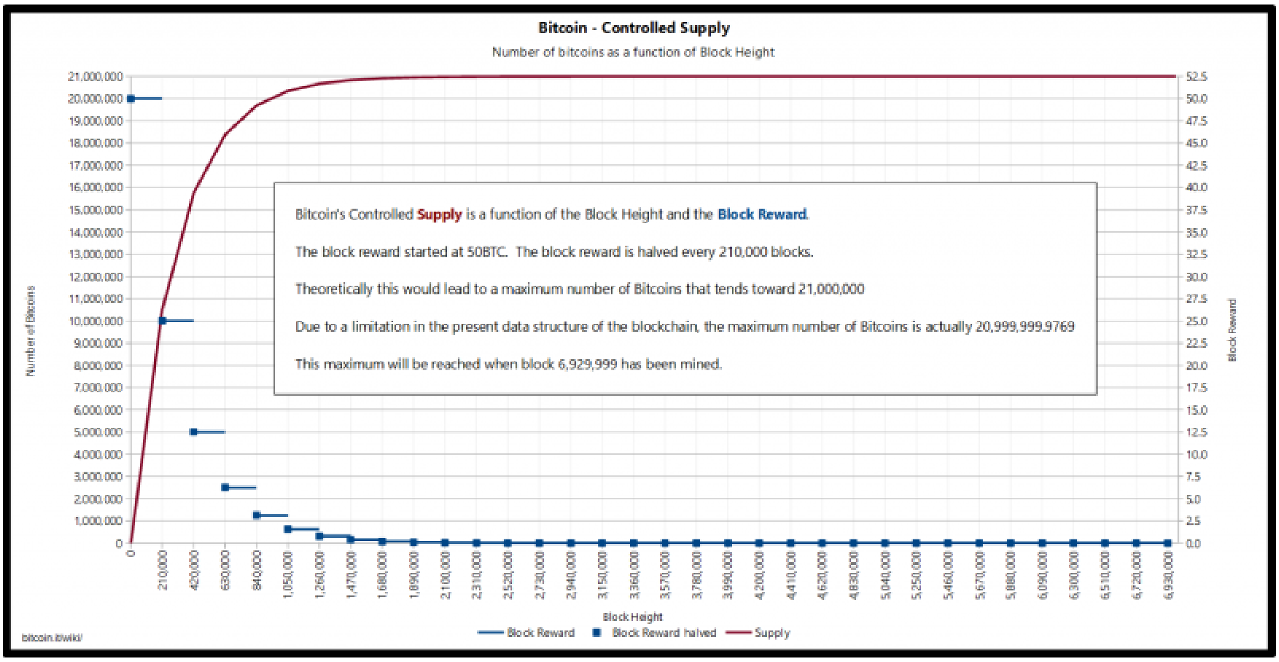
approximately four years.
The result is that the number of bitcoins in existence is not expected to exceed 21 million.

Speculated justifications for the unintuitive value "21 million" are that it matches a 4-year reward halving schedule; or the ultimate total number of Satoshis that will be mined is close to the maximum capacity of a 64-bit floating point number. Satoshi has never really justified or explained many of these constants.

The formula is the following :

$$\frac{\sum_{i=0}^{32} 2100000\,[\frac{50.10^{8}}{2^{i}}]}{10^{8}}$$

This decreasing-supply algorithm was chosen because it approximates the rate at which commodities like gold are mined. Users who use their computers to perform calculations to try and discover a block are thus called Miners.



Bitcoin's Controlled **Supply** is a function of the Block Height and the **Block Reward**.

The block reward started at 50BTC. The block reward is halved every 210,000 blocks.

Theoretically this would lead to a maximum number of Bitcoins that tends toward 21,000,000

Due to a limitation in the present data structure of the blockchain, the maximum number of Bitcoins is actually 20,999,999.9769

This maximum will be reached when block 6,929,999 has been mined.

4.7 Wallet cryptography

A **wallet** is basically the Bitcoin equivalent of a bank account. It allows you to receive bitcoins, store them, and then send them to others.

The name "Bitcoin wallet" is a bit of a misnomer. Bitcoin wallets don't hold actual Bitcoins, those are essentially stored on the blockchain.
Instead, Bitcoin wallets hold the private keys that give users the right to use those

coins. Each Bitcoin wallet comes with at least two keys (multisig wallets can have more) one public, and one private.

The public key lets any Bitcoin user send a sum of Bitcoins directly to any other Bitcoin user, without a middle man. The private key must be kept as secure as possible, since anyone who gets a hold of it has access to every Bitcoin associated with it.

There are several types of wallets out there: Software wallets, web wallets, and paper/cold wallets.

There is a relationship between Wallet and Bitcoin addresses.

Bitcoin address

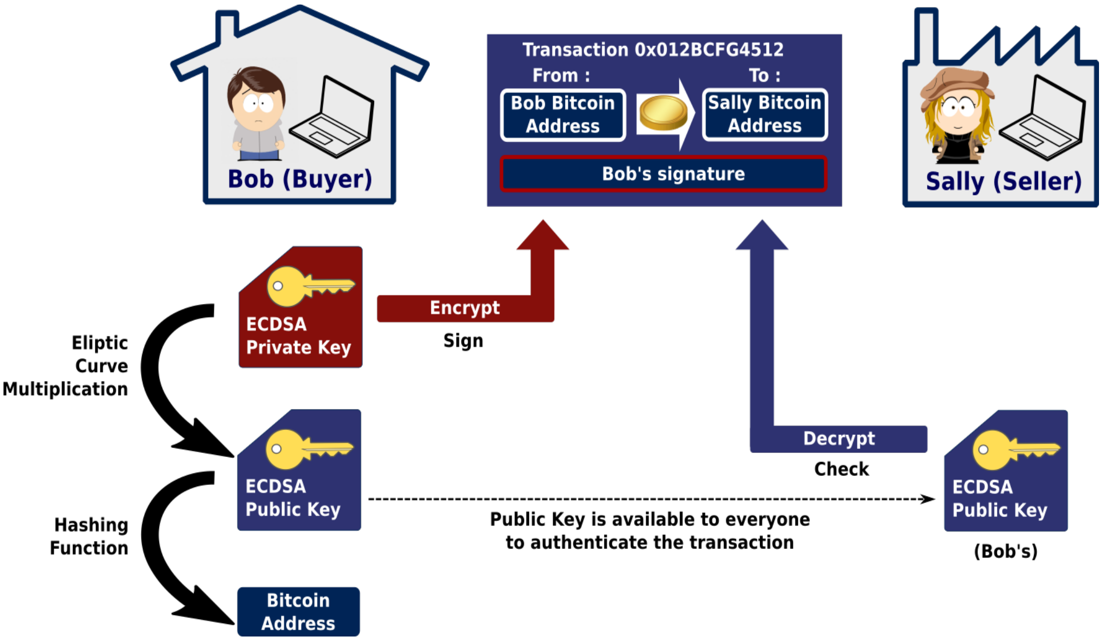
A Bitcoin address, or simply address, is an identifier of 26-35 alphanumeric characters, beginning with the number 1 or 3, that represents a possible destination for a bitcoin payment. Addresses can be generated at no cost by any user of Bitcoin.

There are currently two address formats in common use:

- Common P2PKH which begin with the number 1,
  eg: 1BvBMSEYstWetqTFn5Au4m4GFg7xJaNVN2.
- Newer P2SH type starting with the number 3,
  eg: 3J98t1WpEZ73CNmQviecrnyiWrnqRhWNLy.

A Bitcoin address is a single-use token. Like e-mail addresses, you can send bitcoins to a person by sending bitcoins to one of their addresses. However, unlike e-mail addresses, people have many different Bitcoin addresses and a unique address should be used for each transaction.

A Bitcoin address is a 160-bit hash of the public portion of a public/private ECDSA keypair. Using public-key cryptography, one can "sign" data with your private key and anyone who knows your public key can verify that the signature is valid.



The private key aims at enabling solely and only the owner of an address to create a transaction where the sender of bitcoin is that owner. Keeping this private key secrete is of utmost importance. If someone steals that private key from a wallet owner, he can drag all its money out by creating transaction from that wallet to his own wallet in a perfectly legitimate way.
It's impossible for an attacker to forge a fraudulent transaction from someone else's wallet without having that very private key.
On the other hand, every miner, every network member can safely and fast ensure the validity of the transactions in a block by using the public key of the corresponding wallet owners.

4.8 Merkle Trees

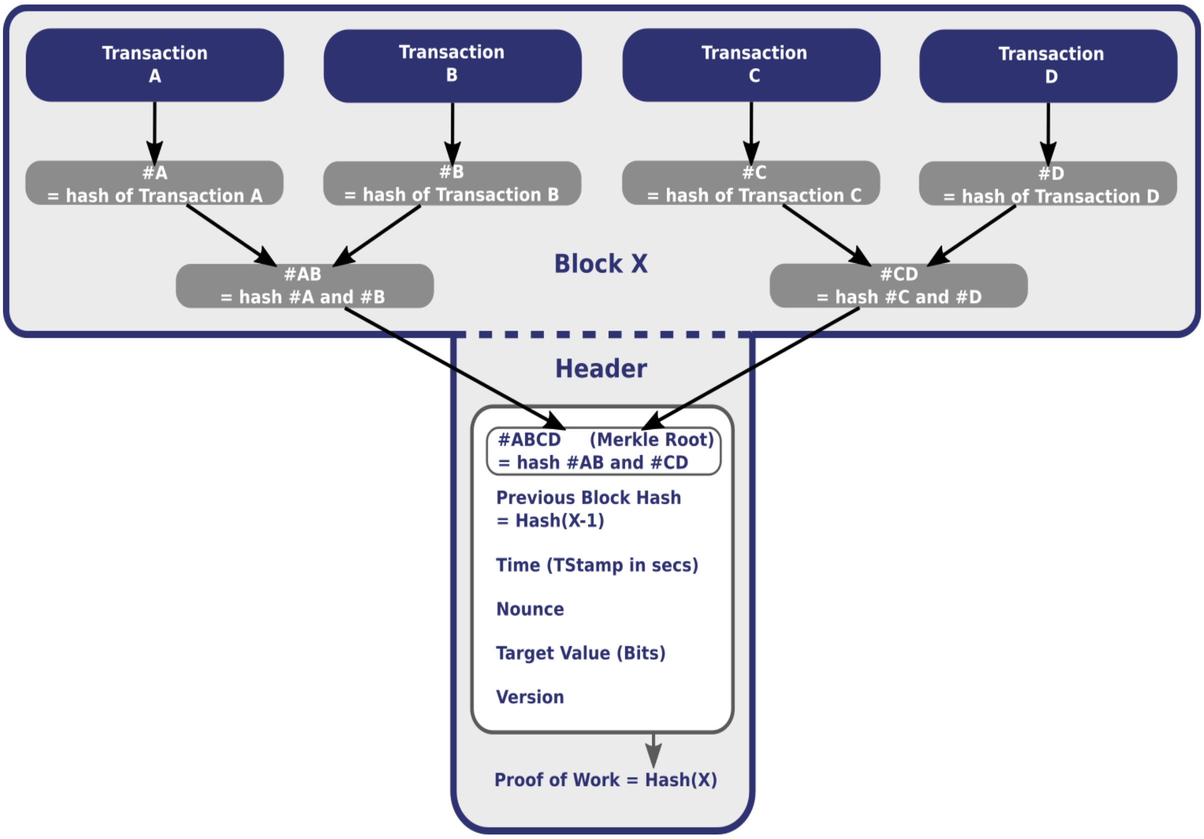
A tree constructed by hashing paired data (the leaves), then pairing and hashing the results until a single hash remains, the merkle root.
The construction of the Merke tree is such that if any single *leaf* transaction is changed, all hashes along the branch would be changed and ultimately the merkle root as well.
This is a key property ensuring security of the blockchain.

Merkle trees in bitcoin use a double SHA-256, the SHA-256 hash of the SHA-256 hash of something.

If, when forming a row in the tree (other than the root of the tree), it would have an odd number of elements, the final double-hash is duplicated to ensure that the row has an even number of hashes.

First form the bottom row of the tree with the ordered double-SHA-256 hashes of the byte streams of the transactions in the block.
Then the row above it consists of half that number of hashes. Each entry is the double-SHA-256 of the 64-byte concatenation of the corresponding two hashes below it in the tree.

This procedure repeats recursively until we reach a row consisting of just a single double-hash. This is the Merkle root of the tree.



In the blockchain, Merke trees enable to verify a transaction in a much faster way than if a validating a transaction meant recomputing the hash from the whole block data.
In addition, as new transactions are broadcasted to the network, updating the merkle root takes only log(n) hash operations on small data instead of re-hashing everything.

Why is the usage of Merkle Tree of so important ?

Think of one thing : the system makes it so that a new block is created every 10 minutes or so. During these 10 minutes, new transactions are taken into consideration immediately by miners. At each and every round, not only the nounce is incremented but the very latest transactions received are added to the block currently being mined.

Without Merkle trees, that would require miners to recompute completely the hash of the entire block, which can be quite a significant amount of work.
Again, with Merkle trees, the number of operations to be computed is log (n) where n is the amount of transactions in the block, i.e. only a few operations to recompute all the hashes along the branches of the new transactions.

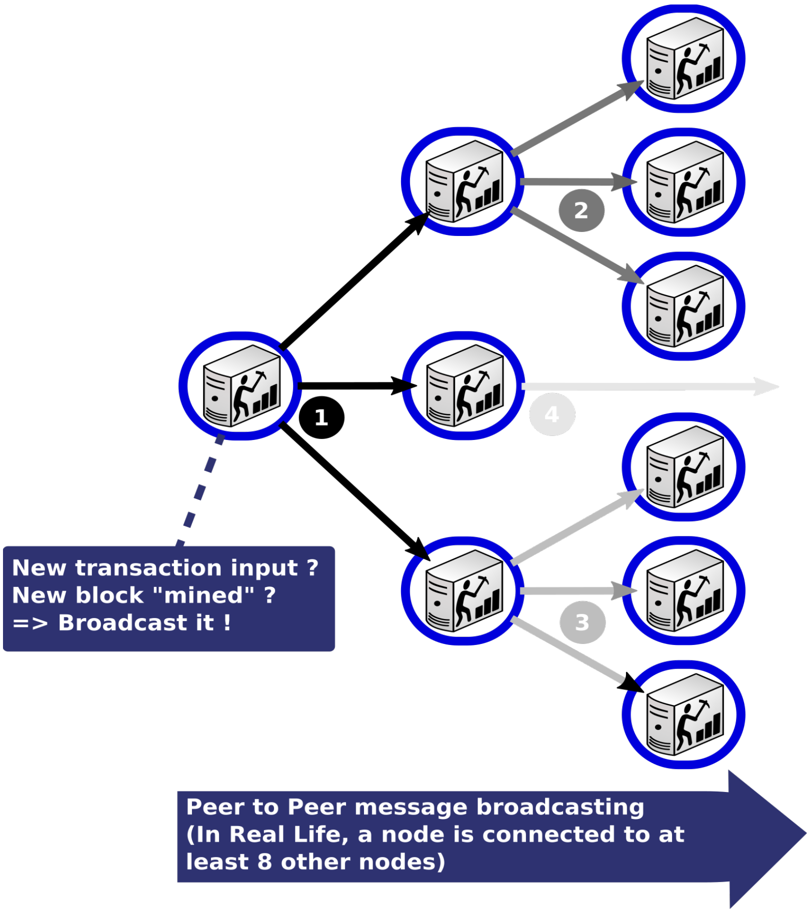
This is pretty important in addition because of the increasing adoption of bitcoin. Since block creation rates is somewhat fixed, the average amount of transactions per block increases making it even more important to avoid recomputing the hash of the full block as transactions are added to it.

4.9 Peer to peer network

A blockchain is a ledger of facts, replicated across several computers assembled in a peer-to-peer network that operates on a cryptographic protocol. In the case of the bitcoin, users send units of currency, by **broadcasting** digitally signed messages to the network using bitcoin wallet software.
In the case of more advanced and recent blockchain technologies - called blockchain 2.0 - (such as Ethereum), users broadcasts Smart Contracts, events or API calls on these Smart Contracts.

Members of the network are anonymous individuals called **nodes**, or "miners" since initially most-if-not-all of them where running the mining algorithm presented above.

Both new transactions and newly mined blocked are broadcasted to the peer-to-peer network.



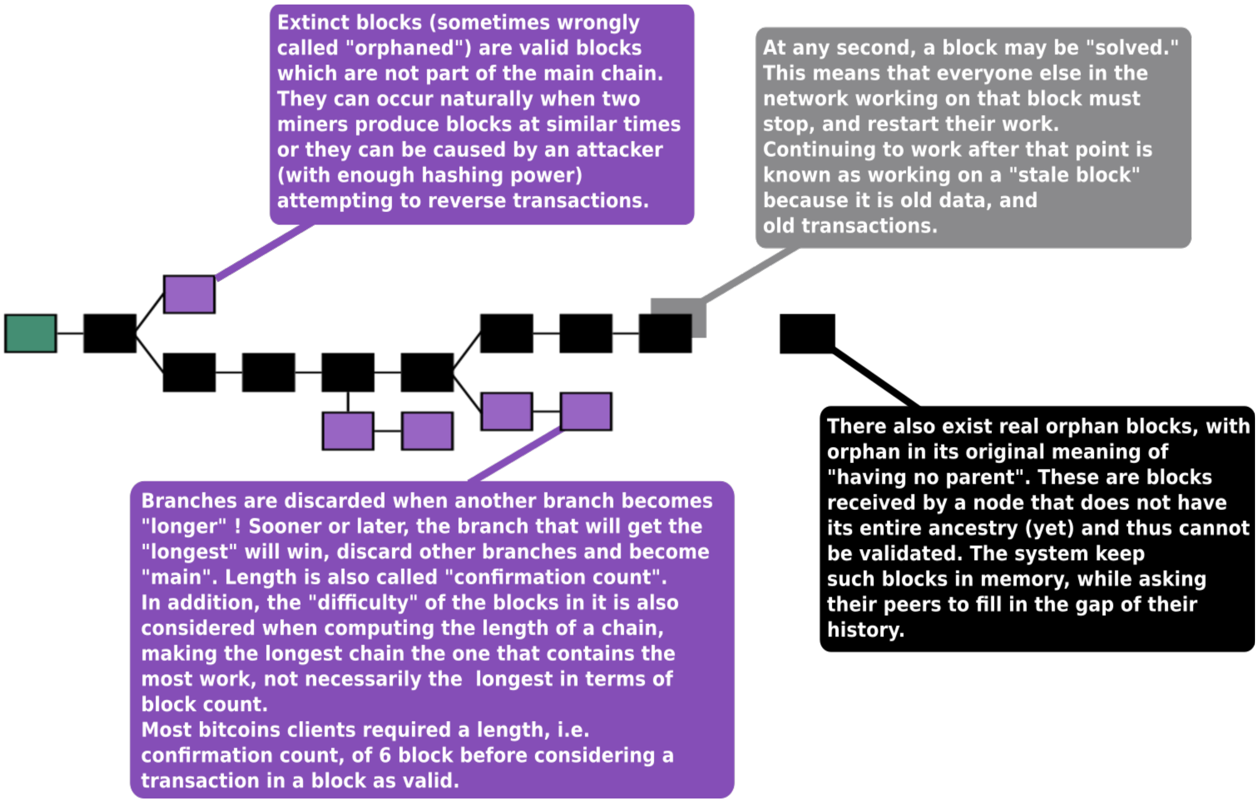
New transaction input ?
New block "mined" ?
=> Broadcast it !

Peer to Peer message broadcasting
(In Real Life, a node is connected to at
least 8 other nodes)

4.10 Orphaned, Extinct and Staled Blocks

Because each block contains a reference to the prior block, the collection of all blocks in existence can be said to form a chain. However, it's possible for the chain to have temporary splits - for example, if two miners arrive at two different valid solutions for the same block at the same time, unbeknownst to one another.
The peer-to-peer network is designed to resolve these splits within a short period of time, so that only one branch of the chain survives.

The client accepts the **longest** chain of blocks as valid. The "length" of the entire block chain refers to the chain with the most combined difficulty, not the one with the most blocks. This prevents someone from forking the chain and creating a large number of low-difficulty blocks, and having it accepted by the network as "longest".



Extinct blocks (sometimes wrongly called "orphaned") are valid blocks which are not part of the main chain. They can occur naturally when two miners produce blocks at similar times or they can be caused by an attacker (with enough hashing power) attempting to reverse transactions.

At any second, a block may be "solved." This means that everyone else in the network working on that block must stop, and restart their work. Continuing to work after that point is known as working on a "stale block" because it is old data, and old transactions.

There also exist real orphan blocks, with orphan in its original meaning of "having no parent". These are blocks received by a node that does not have its entire ancestry (yet) and thus cannot be validated. The system keep such blocks in memory, while asking their peers to fill in the gap of their history.

Branches are discarded when another branch becomes "longer" ! Sooner or later, the branch that will get the "longest" will win, discard other branches and become "main". Length is also called "confirmation count".
In addition, the "difficulty" of the blocks in it is also considered when computing the length of a chain, making the longest chain the one that contains the most work, not necessarily the longest in terms of block count.
Most bitcoins clients required a length, i.e. confirmation count, of 6 block before considering a transaction in a block as valid.

5. Blockchain 2.0

The Blockchain 2.0 is an evolution of the blockchain protocol enabling not only to exchange transaction but rather code and programs in the form of **Smart Contracts**
Now developers are allowed to build programs and API's on the Blockchain Protocol.

This relatively new concept involves the development of programs that can be entrusted with money.
Smart contracts are programs that encode certain conditions and outcomes.
For instance, When a transaction between 2 parties occurs, the program can verify if the product/service has been sent by the supplier. Only after verification is the sum transmitted to the suppliers account.

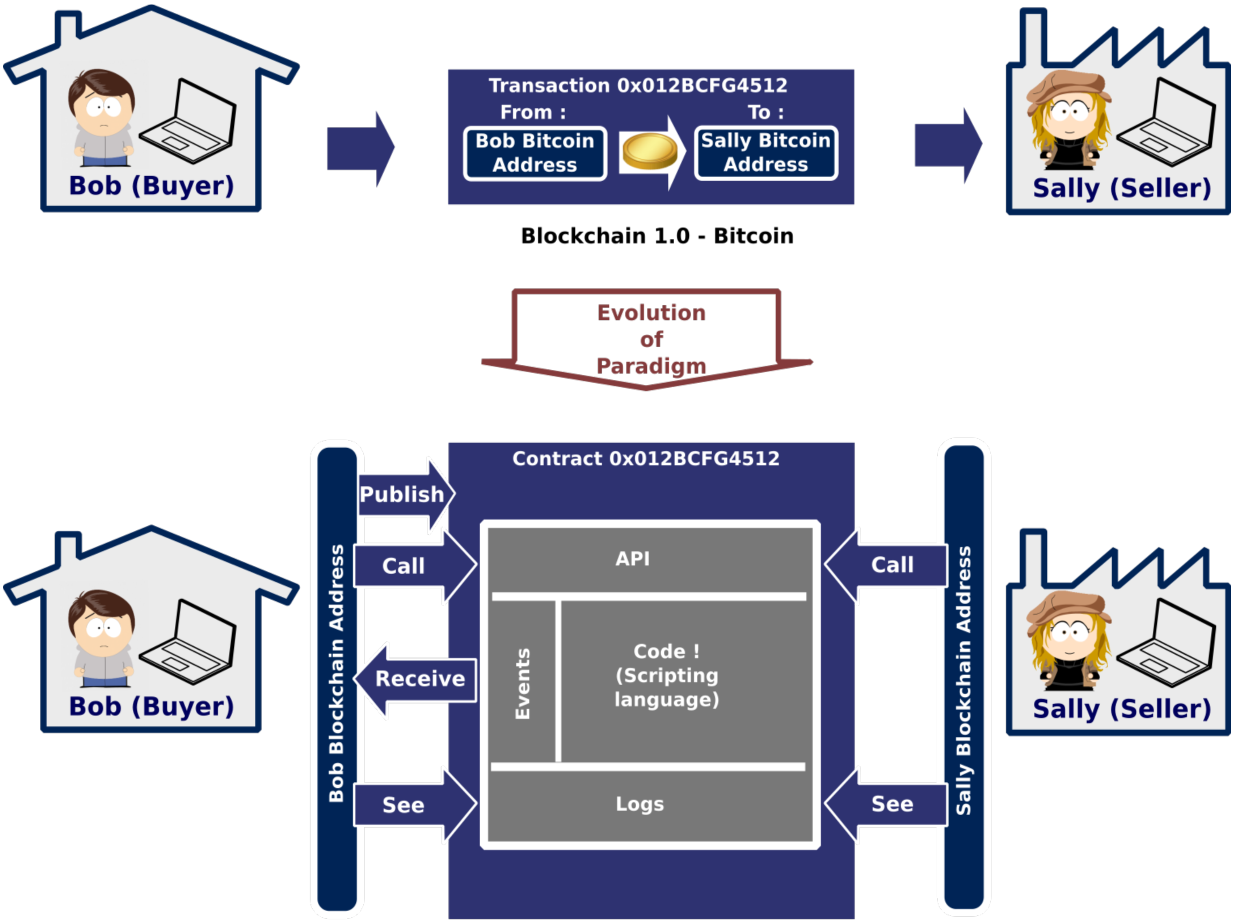
By developing ready to use programs that function on predetermined conditions between the supplier and the client, smart programs ensure a secure escrow service in real time at near zero marginal cost

Apart from Financial transactions, smart contracts are now entering a whole lot of different industry.
For instance in the Legal System, companies like Empowered Law use the public distributed ledger of transactions that makes up the Block Chain to provide Multi-Signature account services for asset protection, estate planning, dispute resolution, leasing and corporate governance.

Ethereum

Ethereum intends to bring together both a crypto ledger and a **Turing-complete programming language**, which is a language can be used to write actual computer programs. They intend to make a browser that is a Swiss-army knife of Block Chain and encryption tools that allow non-technical users to truly leverage the web.

Ethereum aims to implement a **globally decentralized, un-ownable, digital computer** for executing peer-to-peer contracts.
Put more simply, Ethereum is a world computer you can't shut down.



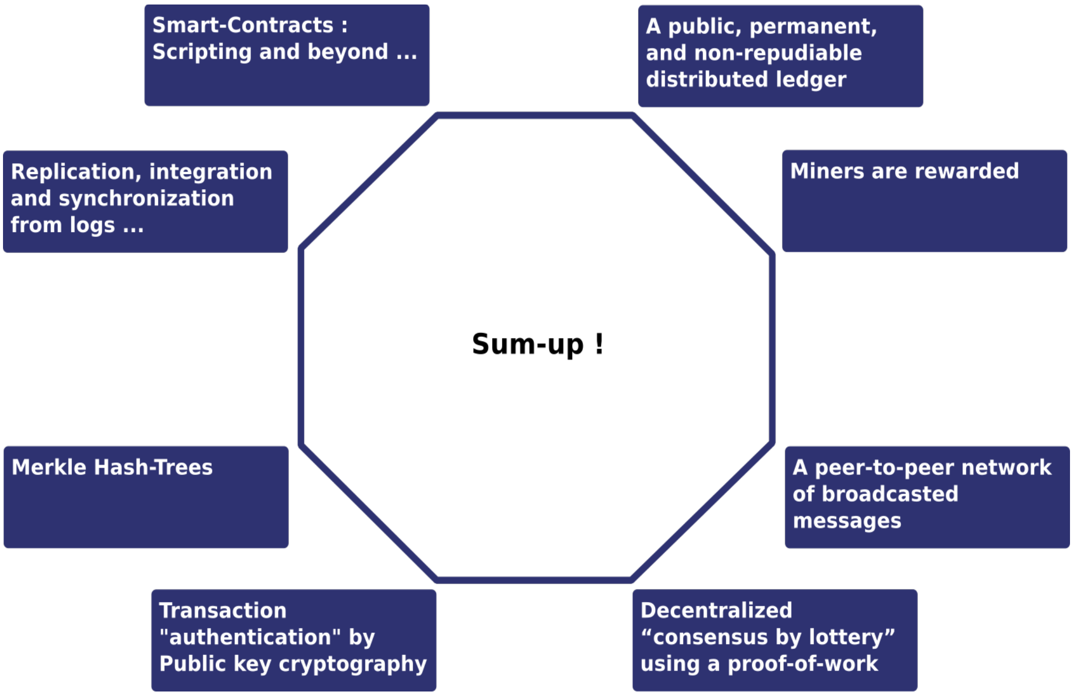
Blockchain 1.0 - Bitcoin

Blockchain 2.0 - e.g. Ethereum, ...

The Blockchain 2.0 and its opportunities will be the topic of a dedicated article expected in the coming weeks on this very blog.

6. Sum-Up

Summing it up, these are the key aspects of the Blockchain technology:


Smart-Contracts : Scripting and beyond ...
A public, permanent, and non-repudiable distributed ledger
Replication, integration and synchronization from logs ...
Miners are rewarded
Sum-up !
Merkle Hash-Trees
A peer-to-peer network of broadcasted messages
Transaction "authentication" by Public key cryptography
Decentralized “consensus by lottery” using a proof-of-work

The Blockchain is an amazing technology. I strongly believe it has the potential to revolution a whole lot of different industries. **That as well will be the topic of a third article on the Blockchain on this blog that should appear in the coming days / weeks.**

One might want to see part of this article as a slideshare presentation available here :http://www.slideshare.net/JrmeKehrli/the-blockchain-the-technology-behind-bitcoin.

Also, have a look at the next article in this serie on this blog : Blockchain 2.0 - From bitcoin transaction to Smart Contract applications