The Blockchain Explained to Web Developers, Part 1: The Theory

François Zaninotto
April 28, 2016

#popular #blockchain

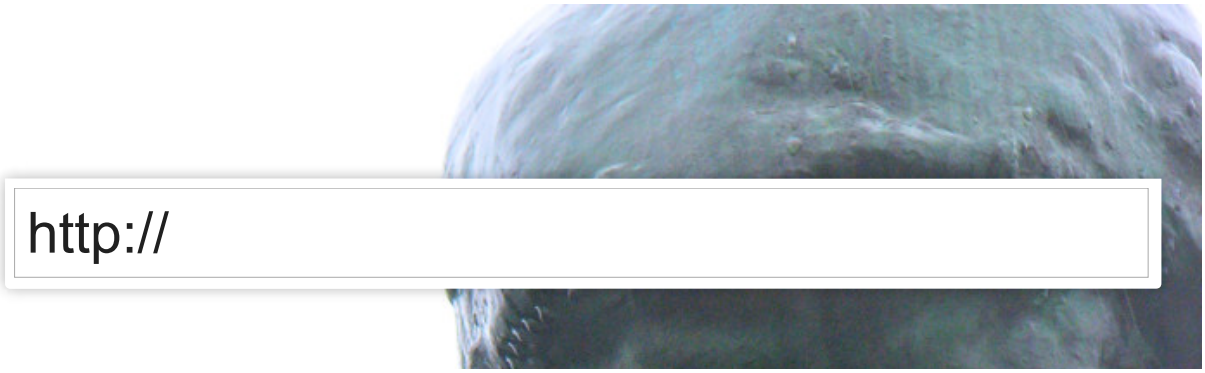
The blockchain is the new hot technology. If you haven't heard about it, you probably know Bitcoin. Well, the blockchain is the underlying technology that powers Bitcoin. Experts say the blockchain will cause a revolution similar to what Internet provoked. But what is it really, and how can it be used to build apps today? This post is the first in a series of three, explaining the blockchain phenomenon to web developers. We'll discuss the theory, show actual code, and share our learnings, based on a real world project.

To begin, let's try to understand what blockchains really are.

What Is A Blockchain, Take One

Although the blockchain was created to support Bitcoin, the blockchain concept can be defined regardless of the Bitcoin ecosystem. The literature usually defines a blockchain as follows:

> *A blockchain is a **ledger** of **facts**, replicated across several computers assembled in a peer-to-peer network. Facts can be anything from monetary transactions to content signature. Members of the network are anonymous individuals called **nodes**. All communication inside the network takes advantage of cryptography to securely identify the sender and the receiver. When a node wants to add a fact to the ledger, a consensus forms in the network to determine where this fact should appear in the ledger; this consensus is called a **block**.*



I don't know about you, but after reading these definitions, I still had troubles figuring out what this is all about. Let's get a bit deeper.

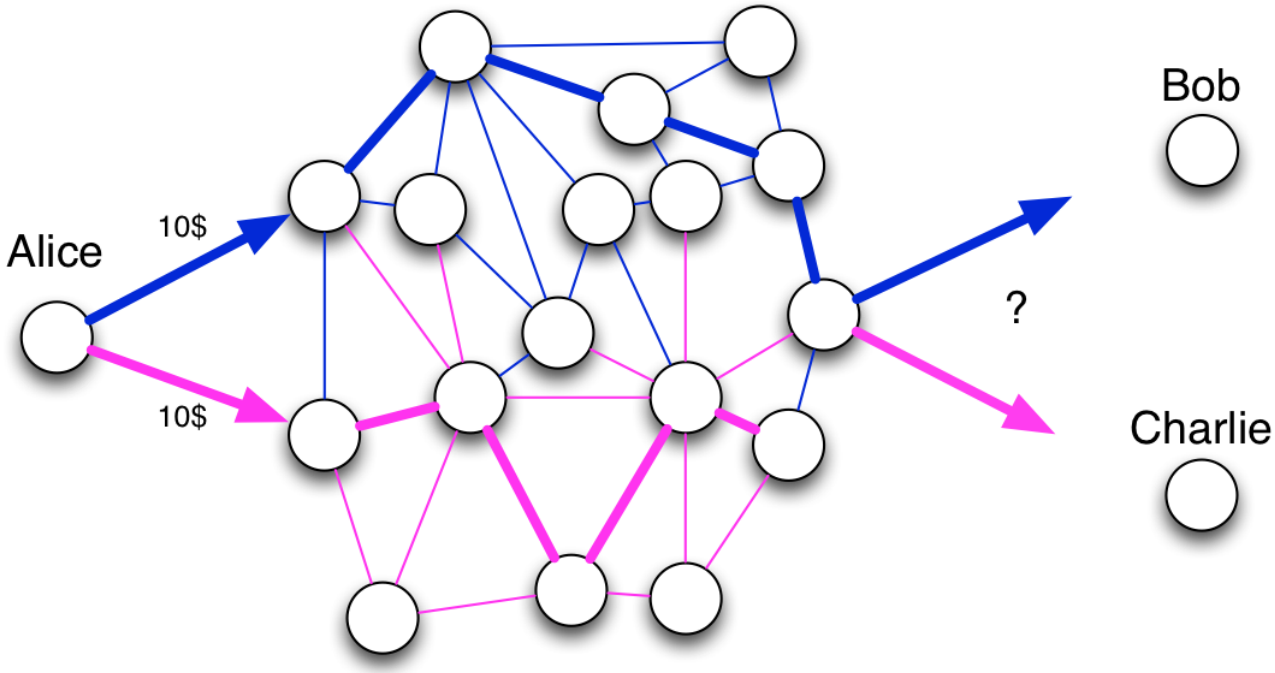
Ordering Facts

Decentralized peer-to-peer networks aren't new. Napster and BitTorrent are P2P networks. Instead of exchanging movies, members of the blockchain network exchange facts. Then what's the real deal about blockchains?

P2P networks, like other distributed systems, have to solve a very difficult computer science problem: the resolution c g in distributed which fact i

http://

> *Take for instance the **double spend problem**: Alice has 10$, and she sends twice 10$ to Bob and Charlie. Who*

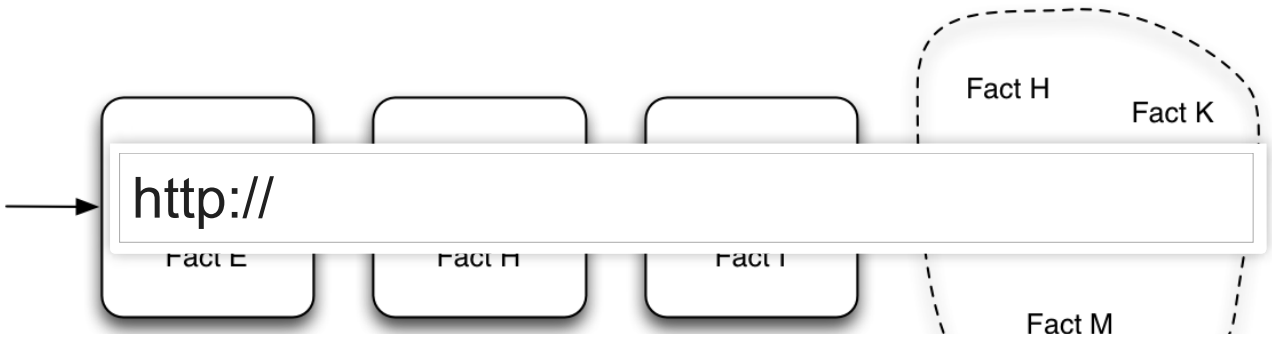


In a P2P network, two facts sent roughly at the same time may arrive in different orders in distant nodes. Then how can the entire network agree on the first fact? To guarantee integrity over a P2P network, you need a way to make everyone agree on the ordering of facts. You need a **consensus system**.

Consensus algorithms for distributed systems are a very active research field. You may have heard of Paxos or Raft algorithms. The blockchain implements another algorithm, the **proof-of-work** consensus, using blocks.

Blocks

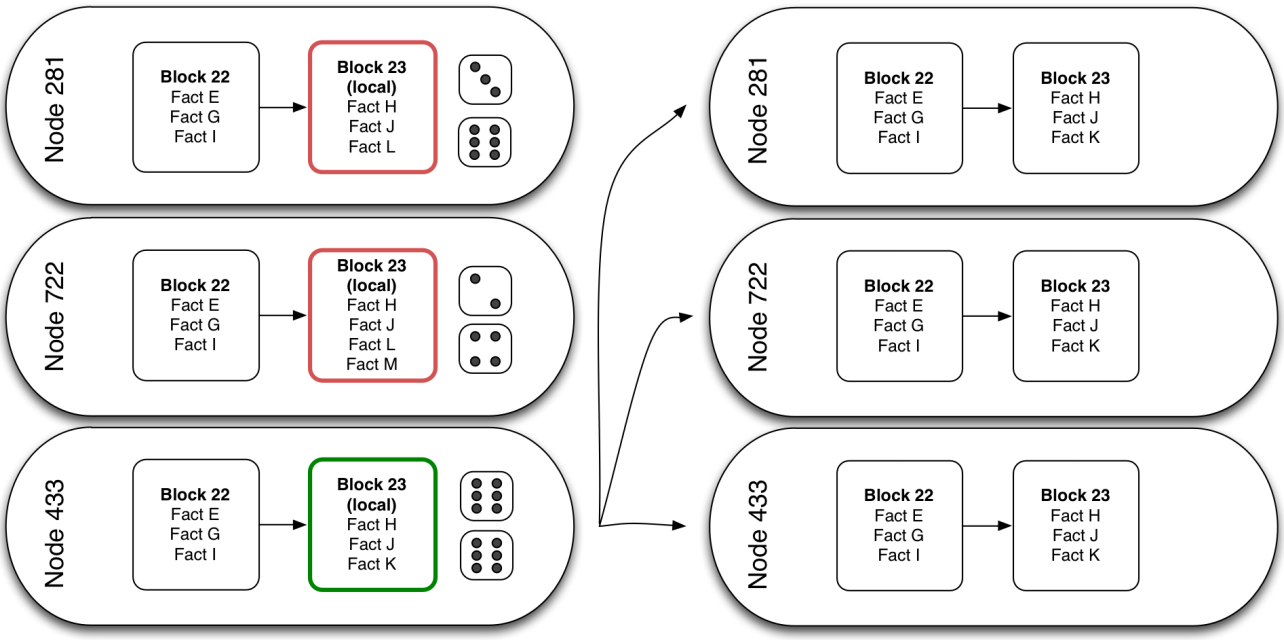
Blocks are a smart trick to order facts in a network of non-trusted peers. The idea is simple: facts are grouped in *blocks*, and there is only a single chain of blocks, replicated in the entire network. Each block references the previous one. So if fact F is in block 21, and fact E is in block 22, then fact E is considered by the entire network to be posterior to fact F. Before being added to a block, facts are *pending*, i.e. unconfirmed.



Confirmed facts Pending facts

Mining

Some nodes in the chain create a new local block with pending facts. They compete to see if their local block is going to become the next block in the chain for the entire network, by rolling dice. If a node makes a double six, then it earns the ability to publish their local block, and all facts in this block become confirmed. This block is sent to all other nodes in the network. All nodes check that the block is correct, add it to their copy of the chain, and try to build a new block with new pending facts.



But nodes don't just roll a couple dice. Blockchain challenges imply rolling a huge number of dice. Finding the random key to validate a block is **very unlikely**, by design. This prevents fraud, and makes the network safe (unless a malicious user owns more than half of the nodes in the network). As a consequence, new blocks gets published to the chain at a fixed time interval. In Bitcoin, blocks are published every 10 minutes on average.

> *In Bitcoin, the challenge involves a double SHA-256 hash of a string made of the pending facts, the identifier of the previous block, and a random string. A node wins if their hash contains at least n leading zeroes.*
>
> ```
> // a losing hash for Bitcoin
> 787308540121f4afd2ff5179898934291105772495275df35f00cc5e44db42dd
>
> // a winning hash for Bitcoin if n=10
> 00000000009f766c17c736169f79cb0c65dd6e07244e9468bc60cde9538b551e
> ```
>
> *Number n is adjusted every once in a while to keep block duration fixed despite variations in the number of nodes. This number is called the **difficulty**. Other blockchain implementations use special hashing techniques that disco*

http://

The process of looking for blocks is called *mining*. This is because, just like gold mining, block mining brings an economical reward - some form of money. That's the reason why people who run nodes in a blockchain are also

Note: By default, a node doesn't mine - it just receives blocks mined by other nodes. It's a voluntary process to turn a node into a miner node.

Money and Cryptocurrencies

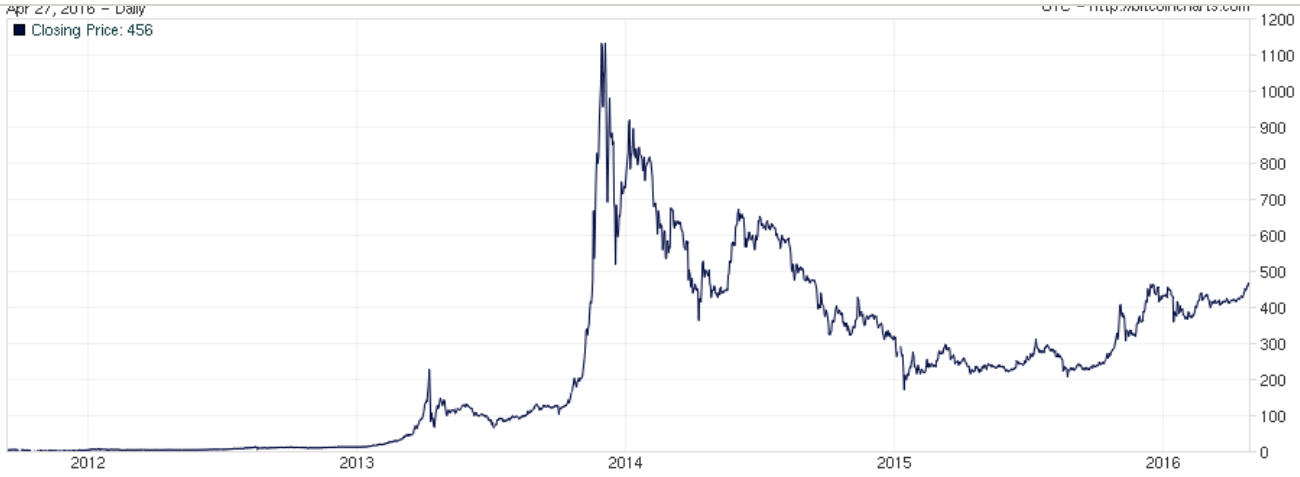
Every second, each miner node in a blockchain tests thousands of random strings to try and form a new block. So running a miner in the blockchain pumps a huge amount of computer resources (storage and CPU). That's why **you must pay to store facts** in a blockchain. Reading facts, on the other hand, is free: you just need to run your own node, and you'll recuperate the entire history of facts issued by all the other nodes. So to summarize:

- Reading data is free
- Adding facts costs a small fee
- Mining a block brings in the money of all the fees of the facts included in the block

We're not talking about real money here. In fact, each blockchain has its own (crypto-)currency. It's called Bitcoin (**BTC**) in the Bitcoin network, Ether (**ETH**) on the Ethereum network, etc. To make a payment in the Bitcoin network, you must pay a small fee in Bitcoins - just like you would pay a fee to a bank. But then, where do the first coins come from?

Miners receive a **gratification** for keeping the network working and safe. Each time they successfully mine a block, they receive a fixed amount of cryptocurrency. In Bitcoin this gratification is 25 BTC per block, in Ethereum it's 5 ETH per block. That way, **the blockchain generates its own money**.

Lastly, cryp
and deman
and hardwa
miners every day, hoping to turn electricity into money. But fluctuations in the BTC value make it less and less profitable.

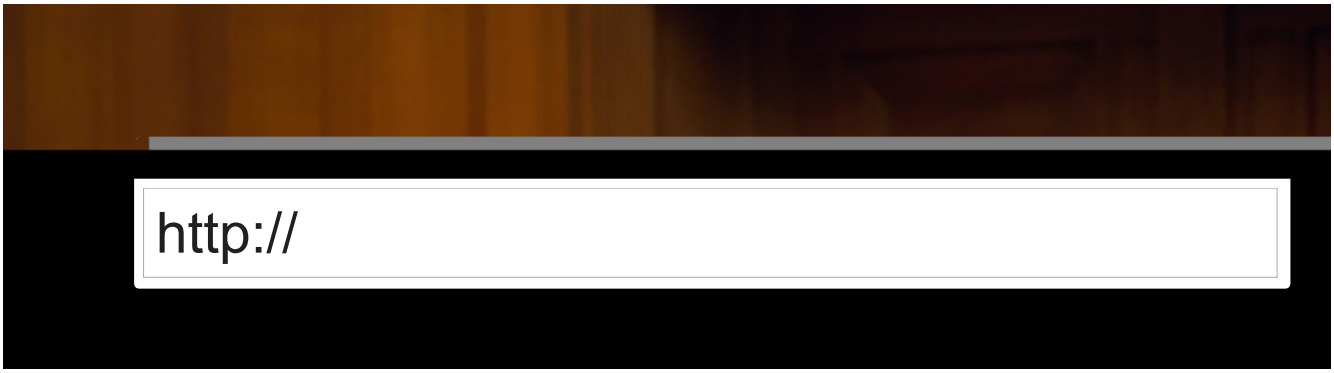


Contracts

So far we've mostly mentioned facts storage, but a blockchain can also **execute programs**. Some blockchains allow each fact to contain a mini program. Such programs are replicated together with the facts, and every node executes them when receiving the facts. In bitcoin, this can be used to make a transaction **conditional**: Bob will receive 100 BTC from Alice if and only if today is February 29th.

Other blockchains allow for more sophisticated contracts. In Ethereum for instance, each contract carries a **mini-database**, and exposes methods to modify the data. As contracts are replicated across all nodes, so are their database. Each time a user calls a method on the contract and therefore updates the underlying data, this command is replicated and replayed by the entire network. This allows for a distributed consensus on the execution of a promise.

This idea of pre-programed conditions, interfaced with the real world, and broadcasted to everyone, is called a **smart contract**. A contract is a promise that signing parties agree to make legally-enforceable. A smart contract is the same, except with the word "technically-" instead of "legally-". This removes the need for a judge, or any authority acknowledged by both parties.

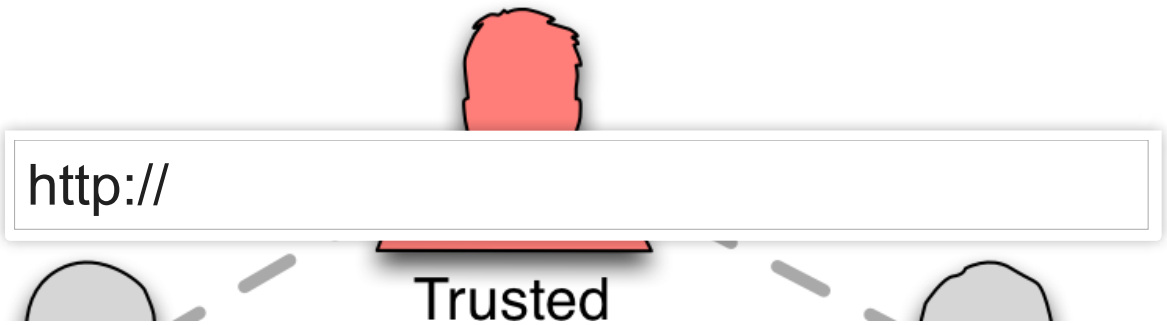


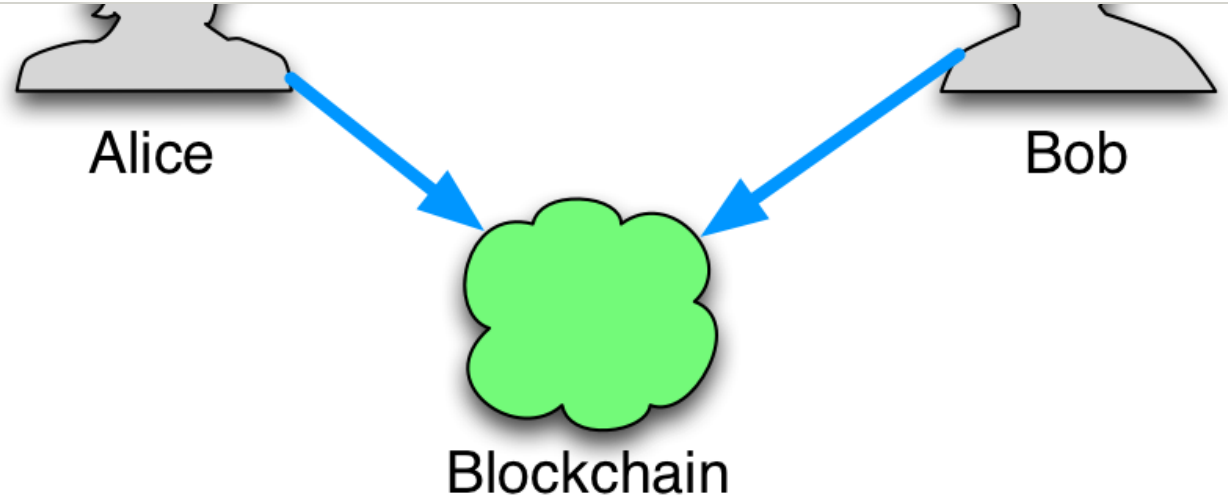
> *Imagine that you want to rent your house for a week and $1,000, with a 50% upfront payment. You and the loaner sign a contract, probably written by a lawyer. You also need a bank to receive the payment. At the beginning of the week, you ask for a $5,000 deposit; the loaner writes a check for it. At the end of the week, the loaner refuses to pay the remaining 50%. You also realize that they broke a window, and that the deposit check refers to an empty account. You'll need a lawyer to help you enforce the rental contract in a court.*
>
> *Smart contracts in a blockchain allow you to get rid of the bank, the lawyer, and the court. Just write a program that defines how much money should be transferred in response to certain conditions:*
>
> - *two weeks before beginning of rental: transfer $500 from loaner to owner*
> - *cancellation by the owner: transfer $500 from owner to loaner*
> - *end of the rental period: transfer $500 from loaner to owner*
> - *proof of physical degradation after the rental period: transfer $5,000 from loaner to owner*
>
> *Upload this smart contract to the blockchain, and you're all set. At the time defined in the contract, the money transfers will occur. And if the owner can bring a predefined proof of physical degradation, they get the $5,000 automatically (without any need for a deposit).*

You might wonder how to build a proof of physical degradation. That's where the **Internet of Things** (IoT) kicks in. In order to interact with the real world, blockchains need sensors and actuators. The Blockchain revolution won't happen unless the IoT revolution comes first.

Such applications relying on smart contracts are called **Decentralized Apps**, or **DApps**.

Smart contracts naturally extend to **smart property**, and a lot more smart things. The thing to remember is that "smart" means "no intermediaries", or "technically-enforced". Blockchains are a new way to disintermediate businesses - just like the Internet disintermediated music distribution.





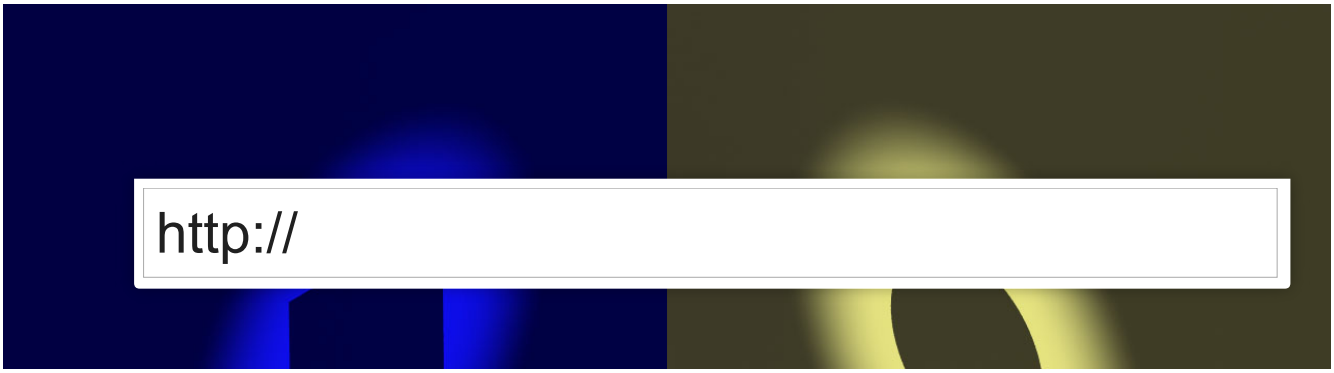
What Is A Blockchain, Take Two

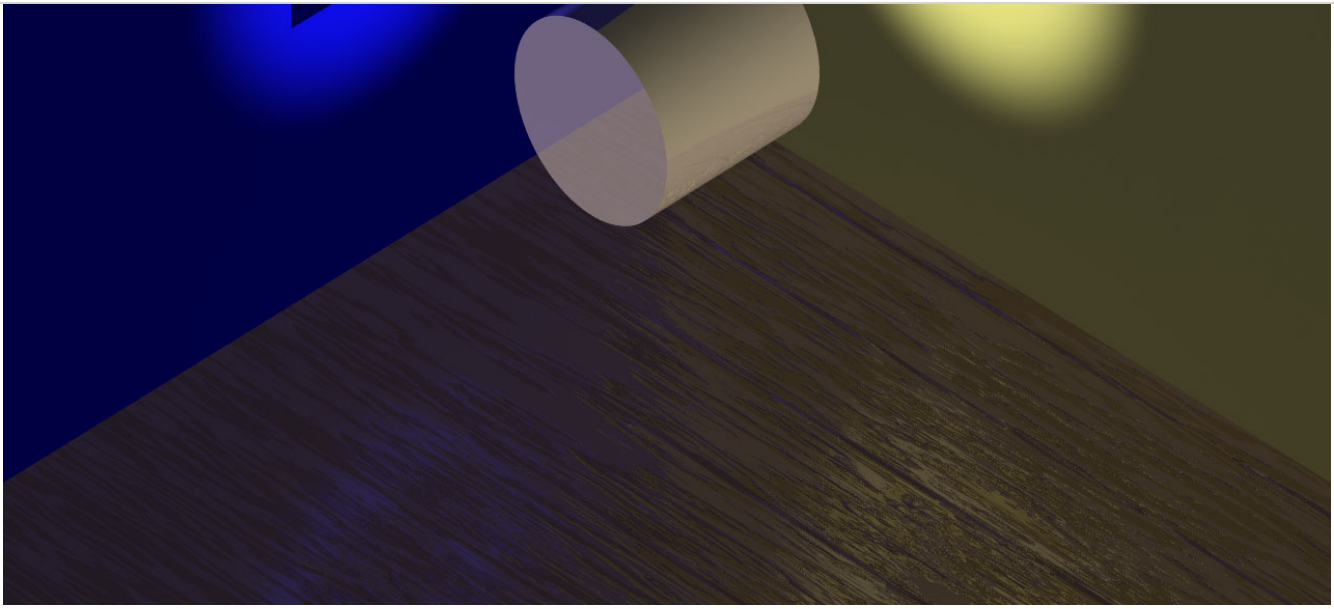
In my opinion, the best way to understand the blockchain is to look at it from various angles.

What it does A blockchain allows to securely share and/or process data between multiple parties over a network on non-trusted peers. Data can be anything, but most interesting uses concern information that currently require a trusted third-party to exchange. Examples include money (requires a bank), a proof or property (requires a lawyer), a loan certificate, etc. In essence, the blockchain removes the need for a trusted third party.

How it works From a technical point of view, the blockchain is an innovation relying on three concepts: peer-to-peer networks, public-key cryptography, and distributed consensus based on the resolution of a random mathematical challenge. None of there concepts are new. It's their combination that allows a breakthrough in computing. If you don't understand it all, don't worry: very few people know enough to be able to develop a blockchain on their own (which is a problem). But not understanding the blockchain doesn't prevent you from using it, just like you can build web apps without knowing about TCP slow start and Certificate Authorities.

What it compares to See the blockchain as a database replicated as many times as there are nodes and (loosely) synchronized, or as a supercomputer formed by the combination of the CPUs/GPUs of all its nodes. You can use this supercomputer to store and process data, just like you would with a remote API. Except you don't need to own the backend, and you can be sure the data is safe and processed properly by the network.





Practical Implications

Facts stored in the blockchain can't be lost. They are there forever, replicated as many times as there are nodes. Even more, the blockchain doesn't simply store a final state, it stores the history of all passed states, so that everyone can check the correctness of the final state by replaying the facts from the beginning.

Facts in the blockchain can be trusted, as they are verified by a technically enforceable consensus. Even if the network contains black sheeps, you can trust its judgement as a whole.

Storing data in the blockchain isn't fast, as it requires a distributed consensus.

Tip

If you have 20 spare minutes to get a deeper understanding, watch this excellent introduction video about Bitcoin, which also explains the blockchain:

How Bitcoin Works Under the Hood

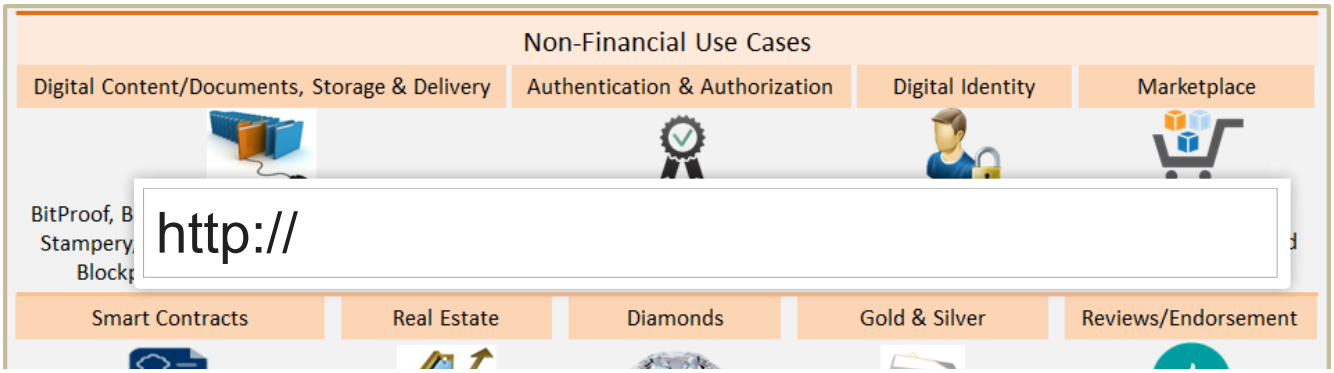
Why It's a Big deal

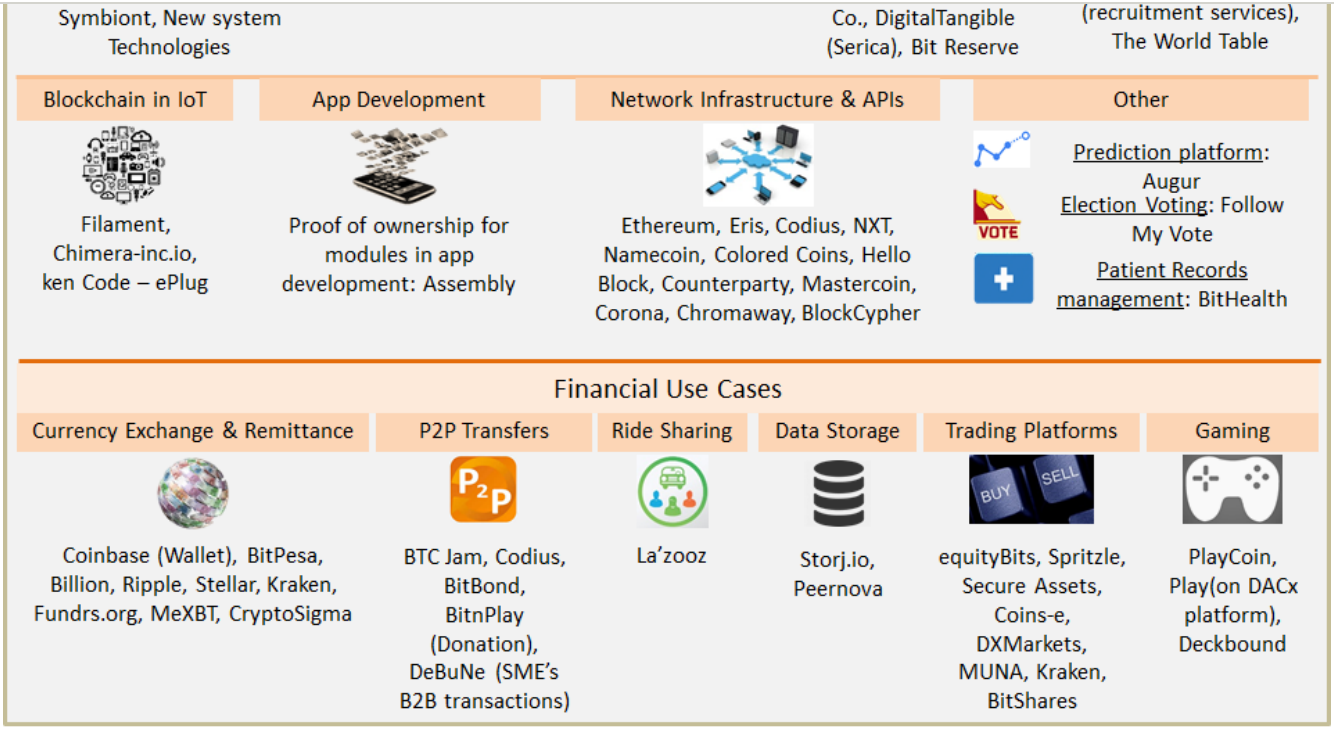
> *«The Bloc*
>
> *«The mos*
>
> *«I think the fact that within the Bitcoin universe an algorithm replaces the functions of [the government] ... is*

http://

These smart people have seen a huge potential in the blockchain. It concerns disintermediation. The blockchain can potentially replace all the intermediaries required to build trust. Let's see a few example applications, most of which are just proof-of-concepts for now:

- Monegraph lets authors claim their work, and set their rules (and fares) for use
- La Zooz is a decentralized Uber. Share your car, find a seat, without Uber taking a fee.
- Augur is an online bookmaker. Bet on outcomes, and get paid.
- Storj.io is a peer-to-peer storage system. Rent your unused disk space, or find ultra cheap online storage.
- Muse is a distributed, open, and transparent database tailored for the music industry
- Ripple enables low cost cross-border payments for banks





Many **successful businesses on the Internet today are intermediaries**. Think about Google for a minute: Google managed to become the intermediary between you and the entire Internet. Think about Amazon: they became the intermediary between sellers and buyers for any type of good. That's why a technology that allows to remove intermediaries can potentially **disrupt the entire Internet**.

Will it benefit to end users, who won't need third parties to exchange goods and services anymore? It's far from certain. Internet had the same promise of heavy disintermediation. Yet Google built the first market capitalization worldwide as an intermediary. That's why it's crucial to invest in the blockchain quickly, because the winners and losers of the next decade are being born right now.

You Won't Build Your Own Blockchain

The technology behind the blockchain uses advanced cryptography, custom network protocols, and performance optimizations. This is all too sophisticated to be redeveloped each time a project needs a blockchain. Fortunately, aside of Bitcoin, there are several open-source blockchain implementations. Here are the most advanced:

- Ethereum: an open-source blockchain platform by the Ethereum Foundation
- Hyperledger: another open-source implementation, this time by the Linux Foundation. The first proposal was published very recently.
- Eris Industries: Tools helping to manipulate Ethereum, Bitcoin or totally independent blockchains, mostly to build blockcha

- Eris for a closed Blockchain, or to discover and play with the technology
- Ethereum for a shared Blockchain

Also, Bitcoin isn't a good choice to build an application upon. It was designed for money transactions and nothing else, although you can program pseudo-smart contracts (but you have to love assembly). The network currently suffers a serious growth crisis, transactions wait in line for up to one hour to get inserted in a block. Miners often select transactions with the highest fees, so money transfers in Bitcoin become more expensive than they are in a Bank. The developer community is at war, and the speculation on the cryptocurrency makes the face value move too much.

Numbers

How big are blockchains today? Let's see some numbers.

Bitcoin:

- Block time: 10 minutes
- Number of bitcoins earned for each mined block: 25
- Number of blocks mined: over 400,000
- Number of transactions per block: over 1,200
- Number of nodes in the network: ~7000
- Bitcoin value: $420
- Most of the computing power is said to be concentrated in China

Ethereum:

- Block time: 10 seconds
- Number of Ether earned for each mined block: 5
- Number of blocks mined: more than 1,400,000
- Number of transactions per day: over 30,000
- Number
- Ether va
- Most of the computing power is said to be concentrated by a miner pool called "Dwarfpool"

Conclusion

The blockchain technology is both intriguing and exciting. Could it be the revolution that gurus predict? Or is it just a speculative bubble based on an impractical idea? After reading a lot on the matter, we couldn't form a definitive opinion.

When we face uncertainty, we know a great way to lift it: trying. That's what we decided to do. Read the next post in this series to see what we've learned by **building a real world app running on the blockchain**.