Blockchains

The good and the bad

Gustav Elmgren

Abstract

A blockchain is simply put a distributed database of digital events or transactions which also contains all past transactions and events that have been verified by the majority of the participants. Blockchain is unique in some ways, because there is no need for third party verifications.

One of the primary purposes of a blockchain is the possibility to be decentralized. Blockchain technology took off in late 2008 when Bitcoin released the white paper “Bitcoin: A Peer-to-Peer Electronic Cash System.” The first transaction was made in January 2009.

This paper will describe how a blockchain works from mainly a Bitcoin, the digital currency, perspective. We’ll talk about blocks and transactions to get a better understanding of how the blockchain work and what possibilities there are. We will also look into if the technology as of today has any drawbacks and how it affects the current state of technology.

At last, we’ll look into some current challenges and benefits that blockchains face.

Table of contents

Research questions

Possible use cases?
Is it useful, if so - what can you use it for?

Challenges
Are there any challenges that could prevent blockchains to be widely adopted?

Benefits
Why is it useful, what are the main perks of using this technology?

Is it currently decentralized?
We will take a closer look if the current state of blockchain is decentralized.

Motive

Because of the characteristics of some Blockchains, more specific smart contracts, it has a wide range of use cases - both financial and non-financial and could have a big impact on those areas.

Method

I have solely used the internet for gathering information. I mainly used google to find sources. Two of the primary sources are white papers from both Bitcoin and Ethereum. I tried to collect as much information as possible from an official site of both blockchains, mainly their GitHub repo and official whitepapers. There is a lot of conferences and lectures about blockchains; therefore, I used YouTube to find and use them.

Criteria

- Reliable publication from a search database
    - Such as Google, Yahoo, or any educational (.edu) publications.
- Publications that seek scientific status
- Availability
    - The publication should be available to everyone.

Literature review

Transactions

Bitcoin, the root of the modern blockchain technology, first appeared in 2008. The whitepaper, “Bitcoin: A Peer-to-Peer Electronic Cash System,” explained a peer to peer cash system that enabled online payments to be transferred directly, without a third party. [1] Because of the impact of Bitcoin [2], I’ll explain what it is and how it works.

Bitcoin is an application that tries to solve the famous problem that is one of the leading issues when talking about digital currency - the double-spending problem - in a decentralized way.
Before Bitcoin, centralized financial institutions severed as a trusted third party who process and validate any electronic transaction.

Every digital transaction made on the blockchain is broadcast to every node in the bitcoin network. Full nodes need to verify two things before recording a transaction:

- Does the person who spends the currency own it?
- Is the balance enough?

A transaction is a transfer of bitcoin value that will be broadcast to the network. It says that X amount of BTC should be removed from your account and be added to the receiver's account. The transaction itself consists of input and outputs.

So what is input and output in a transaction?

Input in a transaction contains a reference to an output from a previous transaction. In other words, if I would send 10 BTC, and got those 10 BTC from two separate transactions, the blockchain would then include both outputs from those two transactions as input, this is called the previous tx. The full input consists of the table below.

Name	Description
Previous tx	Hash of previous transaction(s)
Index	Specific output in the transaction
scriptSig	Contains two components, a signature and a public key of the recipient.

The **output** of a transaction contains instructions for sending bitcoins. It contains:

Name	Description
Value	A number of Bitcoins
scriptPubKey	Contains condition that must be fulfilled to spend those Bitcoins.

When you create a new transaction, that transaction will need, as mentioned before, input which is the output from the transaction that proves you got the BTC in the first place.
This means that you could trace, from the last transaction made, the first transaction ever made. See figure 1 for an overview of input and output in transactions.

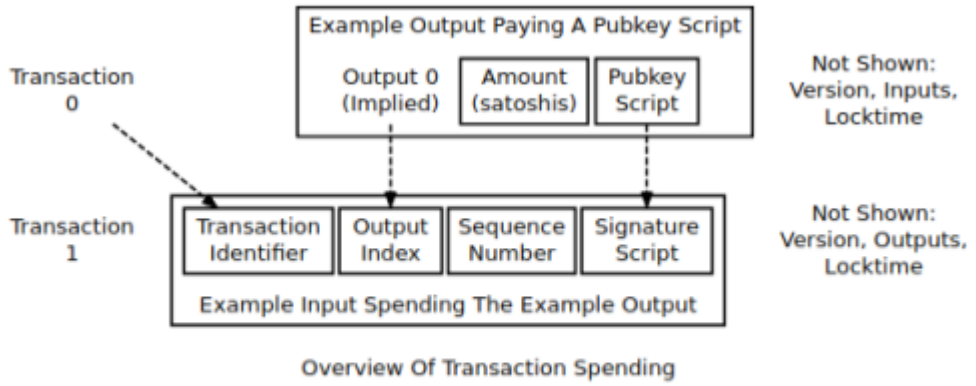


Figure 1: Overview of Transaction Spending (from https://bitcoin.org/en/developer-guide#transactions)

If Alice makes a transaction to Bob. How do we know Alice created that transaction, and that Bob received them? Public-key cryptography is perfect for this. Basically data encrypted with a public key can only be decrypted by using a private key data signed with the private-key can be verified by using the public-key.

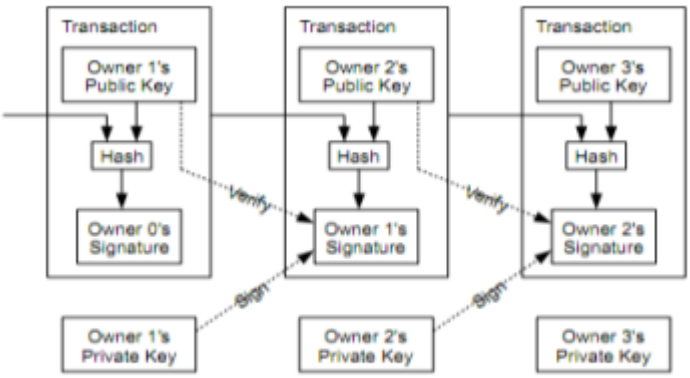


Figure 2: The chain of ownership (from https://bitcoin.org/bitcoin.pdf)

The signature in the new transaction creates a verifiable link between the new transaction and the old one, see figure 2. The new transaction points to the old one explicitly, and the new transaction's signature can only be generated by the holder of the private-key of the old transaction (the old transaction explicitly tells us who this is through the owner-pubkey field). So the

old transaction holds the public-key of the one who can spend it, and the new transaction holds the public-key of the one who received it, along with the signature created with the spender's private-key.

Blocks

Every transaction gets placed into a so-called block, that block consists of several transactions. A transaction is not broadcast and put in a block in the order they were generated (therefore the problem with double-spending still exists). Every block (except the first one, this is called the genesis block and will not be discussed here) is pointing to the hash in the previous block, hence the name blockchain. See figure 3.

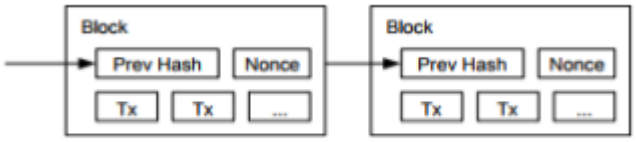


Figure 3: from (https://bitcoin.org/bitcoin.pdf)

The point of a blockchain is you need to reach an agreement between all parties, a so-called consensus. This is vital because no central party could verify if a transaction is valid or not. So how do you make everyone to agree that Alice did, in fact, send Bob Bitcoins and as important - be sure that no one changes it after the transaction was confirmed?

Proof of work, a consensus algorithm, is something Bitcoin uses. The transaction is made and put on a block, but the block does not exist on the blockchain - yet.

First, let's shortly talk about hash functions. If I give "abc" as input to a function, I would, for example, get "cbd1" as output. It is tough to know that "cdb1" (output) came from "abc" (input). Every time "abc" is used as input, "cdb1" will be the output. One possibility to guess the input from a given output is to try every input we can test, starting at 0. Eventually, we will guess right. How could someone else verify it? By testing the given input and see if it matches the known output.

So, in other words, a given output is complicated to calculate the input. But given both input and output, it is pretty easy to verify if the input leads to the output.

To create a block, an output is given, and every full node has to solve the puzzle by providing work (CPU power) otherwise called as **Proof Of Work**. As you can see in figure 4 there is something called a nonce, all you need to know about a nonce is it's just a 'meaningless' number that could be changed as many times as you like to generate different hashes. You have the possibility to change this because there is a certain target that the hash must follow. If the hash (output) is incorrect, you can change the nonce to get it right. An example of a target could be must start with 0000.

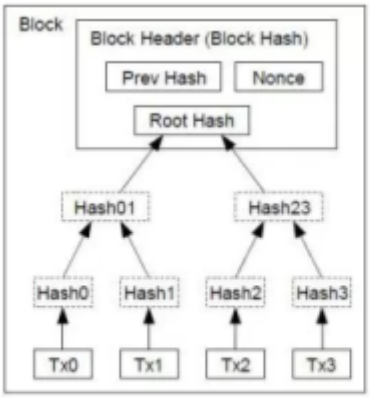


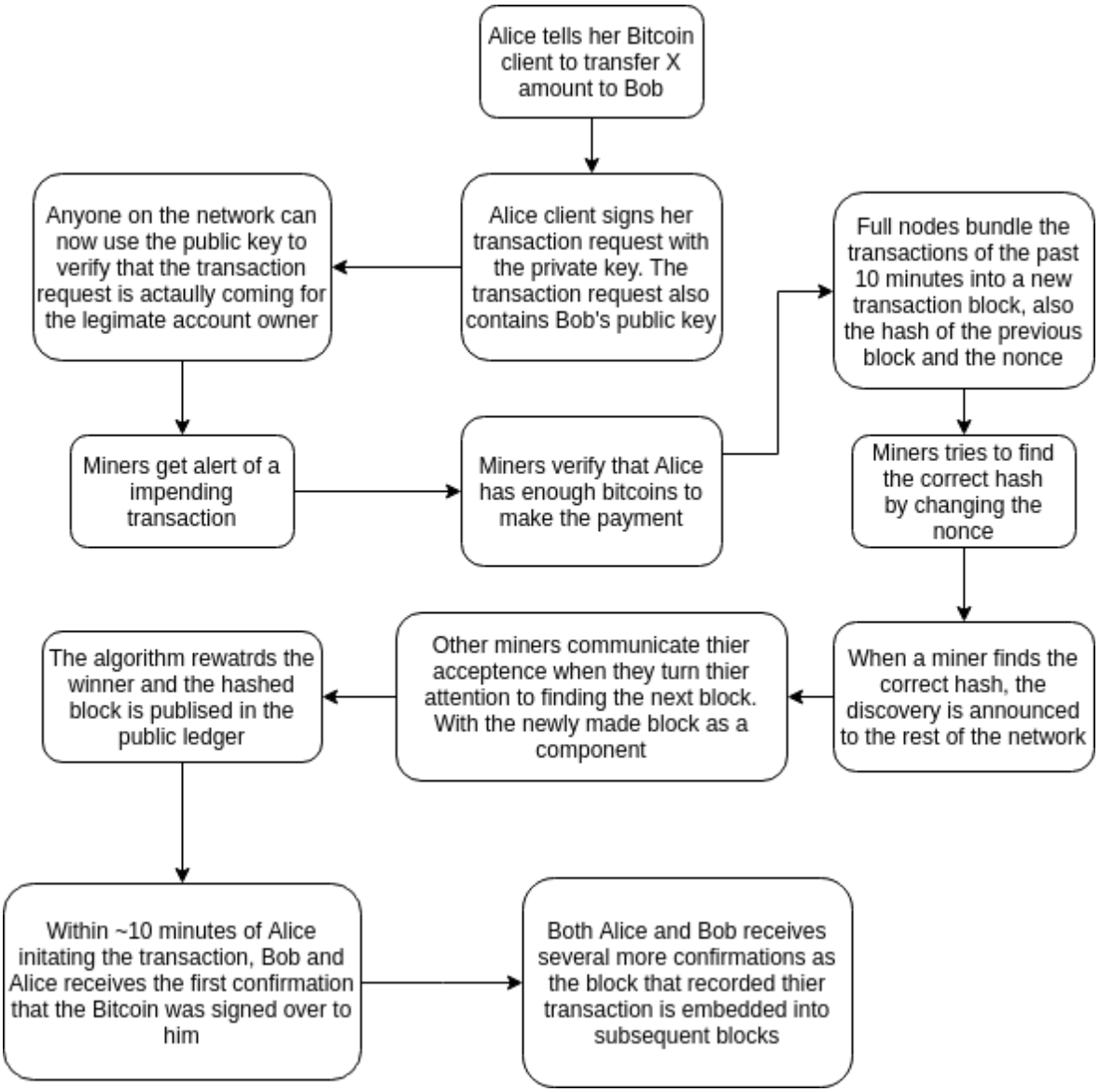
Figure 4: Merkle tree. (from: https://bitcoin.org/en/developer-guide#block-chain)

All full nodes need to hash the previous block, a nonce, root hash (and a few more things that won't be discussed); this is the

input to get a hash (the output) that satisfies the target hash.

When a block is solved by a miner node, it announces it to all other nodes. The other nodes could easily verify it by having the output and the input. But if someone doesn't agree with the node that announced the solution? Not much, the majority of people decides what block should be added in the blockchain. All the other nodes have to be forgiving and accept the longest blockchain.

We talked separately about transactions and blocks. Below is a figure that explains the flow of when Alice creates a transaction to Bob and ends with Bob receiving the sent Bitcoins.



Consensus algorithms

As we now know, both Bitcoin and Ethereum use Proof Of Work at the moment to prevent double spending. It also rewards those who solve the task by creating and giving the digital currencies. But it seems to have a few drawbacks.

There is something called Proof of Stake, another consensus algorithm. Instead of proving how fast you can calculate the puzzle. You need to prove how much, in this example, Ether you would risk verifying that block. The rewards are distributed according to how much proof of stake you have. [13]

One benefit using proof of stake is that you save a lot of computational power. It's estimated Bitcoin and Ethereum burn over 1$ million worth of electricity. Another, more financial aspect, is that you need to lock up Ethereum which means the price should go up. There is also the ability to use economic penalties to make various forms of 51% attacks vastly more expensive.

It's important to clarify, in the case if a 51% attack, the attacker cannot:

- Change the number of coins generated per block
- Create coins out of thin air
- Send coins that never belonged to him

But it's possible to:

- Prevent transactions from gaining any confirmations
- Prevent all other miners from mining any valid blocks
- Double-spend transactions

Back to Proof of Stake. Every node places a bet on a block. The block still needs to be honest. A node, at random, based on the amount of X placed, get picked. If a node put a bet on a dishonest he/she will get penalized. The minimum bet on Casper is 1250 ETH. 1250 ETH is roughly worth $500 000. Only a small percent of the people have that kind of volume. Over a period, people will get way richer, and the network will lie in the hands of a few people. That means a few people will have control over the network.

It's important to point out that PoW and PoS are not the only consensus algorithms. And there are different implementations of each one of them. Except for PoW and PoS there is also Proof of Publication, Proof of Burn and Proof of Ownership. These will not be discussed in this paper.

Analysis and discussion

Possible use cases

Blockchains has many other use cases than a raw currency; there is about 900 applications today that use the blockchain technology (some identical, some not). Although the concept of the blockchain was born out of research into cryptocurrencies.

"[Blockchain] is to Bitcoin, what the internet is to email. A big electronic system, on top of which you can build applications. Currency is just one." - Sally Davies

Ethereum, another blockchain application, is not just a digital currency like Bitcoin. It has a feature called smart contracts. A smart contract is a piece of code that can execute itself at a certain time. When you create a transaction with Bitcoin, the blockchain executes a script ("pay to pubkey hash"). [3]

```
<sig> <pubKey> OP_DUP OP_HASH160 <pubKeyHash> OP_EQUALVERIFY OP_CHECKSIG
```

The key feature in Ethereum is the element of smart contracts; any algorithm can be encoded in these contracts. They are run by each node as part of the blockchain creation process. They even have an own address in the blockchain. In other words, the smart contract is not carried inside each transaction that makes use of it.

A node can create a special transaction that assigns an address to a contract. This transaction can also run code at the moment of creation. After the creation, the contract becomes forever a part of the blockchain and its address never changes. Any node can call any of the methods defined by the smart contact by sending a message to the address for the contract and specifying data as input and the method that should be called.

But Ether is a currency, just like Bitcoin? No, Ether is more like an asset.

As mentioned before, it's possible to call a method in a smart contract. It's a computation that runs as part of a transaction. Furthermore, this means each node must be able to run computations. But what if that code would run forever?

This poses a big problem for Ethereum, no node can get caught up in an infinite loop running a single program. This would eventually halt all transactions. Here is when Ether comes into the picture. Every time a script is run, the user requesting the script must set a limit of Ether to spend in it. So in other words - it cost Ether to run a script - thus making it costly to have a script running indefinitely.

Big companies like IBM [4], Samsung [5] and Amazon [6] are all companies that work with alternative uses of the blockchain for their applications. Even banks are interested, nine of the biggest banks join to form blockchain partnership. [7]

When we now have a better understanding the possibility of smart contracts we might see other use cases, some of which are digital identity and voting. [8] [9] [10]

Personal identifications, such as passports could be stored at the ledger. It could be stamped with a private and public key just like a Bitcoin wallet which only requires the private key to unlock it. Because of the level of transparency digital voting could be huge. You could, for yourself, see if your vote was successfully transmitted while remaining anonymous to the rest of the world.

Is Bitcoin decentralized?

We talked about how Proof Of Work solved the double-spending problem. But that solves the problem **only if** the computing power is well **distributed**. If we look at the hashrate distribution [11] we can see that four of the most significant mining pools control more than 51%. So it's safe to say that this processing is getting more centralized. If they were to group, they could prevent transactions from being executed, approving a specific set of transaction and double-spending transactions.

There is also the problem with activities such as updating protocol and solve incidents that could occur within the code. These activities are not, per design, decentralized and a few people at Bitcoin Foundation is controlling this. But from another aspect, every user can choose to use what version they like best. And for the blockchain to work, a majority of people need the same version. So it's in the developer's best interest to make the users happy. [12]

We know that miners work on extending the longest blockchain. It's also possible to work on different versions of the blockchain. Blockchain forks are detrimental to the operation of the Bitcoin system. Since one blockchain will eventually prevail (the longest), all transactions that were included in all other chains will be invalidated by the miners in the system. Bitcoin has no solution to this. But if a fork persists for a certain time, the developers need to take action to favor on a chain or the other. This could be done by hard-coding the preferred chain in the client code.

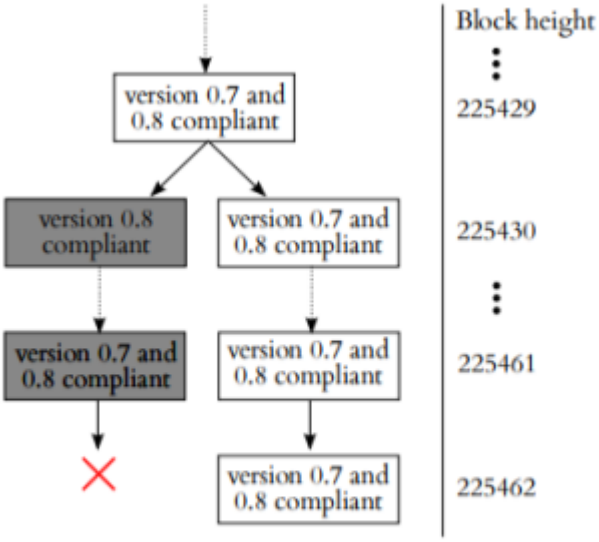


Figure 5: Sketch of the block chain fork that occurred. (from: https://eprint.iacr.org/2013/829.pdf)

In March 2013 a chain fork happened (updated from 0.7 => 0.8, see figure 5). The chain adopted by version 0.8 clients was supported by the majority of the computing power in the network. Nevertheless, the Bitcoin developers decided, 90 minutes after the fork occurred, to force the smallest chain to be the "genuine" one, and convinced the owner of the biggest mining pool, Eleuthria from BTC Guild, to support this decision.

This decision comes at odds with the claim that Bitcoin is a decentralized system and that the majority of the computing power regulates Bitcoin. Less than ten entities decided to outvote the majority of the computing power in the network.

Vitalik also point out that blockchains today does not, necessarily, protect against common node failure. [27]

Some problems may be:

- All nodes run the same software client, and the client has a bug
- All nodes run the same software client, and the developer team is corrupt
- The majority of mining hardware is built by the same company, which could contain a backdoor

We could see that all process of this technology for now (verifying blocks and maintaining the code) is not as decentralized as it could be in the future.

Challenges

Scaling

One of the most significant problems is how current applications scale. [14] At this time, every participating node in the network must process every transaction and maintains a copy of the entire state. This is because it's decentralized in some sense. Bitcoin's current blockchain size is about 140 GB. [29]

If Bitcoin were to process the same amount as Visa does (~2000 transactions per second), it would grow by 1GB per **hour** (in September 2017 it grew by 6 GB). Ethereum will most likely suffer a similar problem. [13]

The maximum throughput (transactions per second) of Bitcoin and Ethereum is somewhat different though. [25]

- Bitcoin: Maximum block size divided by block interval. (~3-7 tx/s)
- Ethereum: Maximum gas limit divided by block interval. (~5-15 tx/s)

One important thing here is Ethereum gas limit is not hard-coded as Bitcoin's block size.

If the blockchain would be a relatively large size (50TB), it would likely be big companies that would run full nodes (mining nodes). The clients would then connect to "SPV" nodes. SPV stands for Simplified Payment Verification, and that means the client does not verify everything but relies on connecting to a trusted node (full nodes). This means it would be rather more centralized that only a handful few with a high amount of processing power would confirm all transactions. [13]

There are a few proposals on how to solve this. Some of which are SegWit (Bitcoin), increase blocksize (Bitcoin), Off-chain state channels [15], Sharding [16], Plasma and off-chain computations.

Most likely, any solution that involves increasing block size won't work - it will only advance the problem for a short period. [17]

> "I don't think we have a good solution because the technology fundamentally doesn't scale yet." - Peter Todd, Bitcoin Core Developer [17]

Government regulations

In the world of crypto there is something called ICO. It stands for **I**nitial **C**oin **O**ffering. It's more or less a company that offers their currency or asset to the people for a limited supply and time. This is rather risky and open for everyone, with no protection for the consumer. There is also an argument that virtual currencies are "tools for criminal activities of money laundering, drugs deals, smuggling and illegal fundraising." Because of this, some countries see cryptocurrency as a threat and fights against them. [18]

Adoption

There is a long way for this technology to be widely adopted. Educations is a rather big problem for the adoption of the blockchain. [19] People who don't understand the value won't buy in or use it. You have to consider who needs to use this in order for it to scale. It's not just technicians and engineers. It's everyone from CEOs to heads of marketing.

It's important to say that everyone does not need this knowledge. There is plenty of end users who do not now know how the internet works but use it broadly.

We also know that today, the technology can't handle massive adoptions because of how bad it scales. So it could be somewhat good that it will take a while for it to be adopted. Time and adoption is not anything new. One example is TCP/IP, the communication protocol used on the internet; and it took more than 30 years for TCP/IP to move through all the phases of reshaping the economy. [20]

Benefits

As pointed out above, blockchain is still in an early stage and have some big obstacles on the way to be widely adopted. But it still has some rather big advantages.

Transparency

One beauty of the blockchain is almost everything is open source [21] [22] [23], which means you can see exactly how the transaction will be executed and get a good overview of the source code. This requires some technical knowledge, but as soon someone sees something that would have a negative impact they would start spreading that information on various platforms.

Recognition at universal level

A blockchain is not tied to any country or corporation [24]. Because of this, a transaction can be made from anyone, including anything digital of value, to anyone. There is no limitation location wise.

User control

You have the possibility to be your own bank. It's not necessary though, but rather recommended. It's a major benefit from this technology. What I mean by this is you own the private key, that is used in the signature of a transaction to create one. It's only you who can create that signature. Being your own bank creates responsibilities. If you lose it, the wallet is forever gone.

Decentralized

In earlier parts of the report we talked about how some process isn't decentralized (development, and in some cases PoW - also the hardfork in March 2013 regarding Bitcoin). But some parts are in matter of fact decentralized. Blockchains are architecturally decentralized (no infrastructural central point of failure). And in some cases, you could also claim they are politically decentralized. [27] [28]

Immutability

Once a transaction is validated and recorded and for every block after that, it gets more and more difficult to change any data. You would also, over time, get an immutable audit trail which makes it possible to track transactions. [30] This is not necessarily true on all blockchains, like Monero that aims to have full privacy. [31]

Eliminations of intermediaries

Because the blockchain achieves consensus on the network there is no need for third-party intermediaries to verify or transfer ownership. This could lower both cost and processing time because no person/cooperation needs to spend the time to verify .i.e does not cost money.

Conclusion

Can a blockchain remain genuinely decentralized - not just in theory but practice? I don't think so. There is no right answer how to keep the chain decentralized, yet. However, I believe better algorithms and methods will evolve. I would say blockchains are decentralized from an architectural aspect, but people tend to exaggerate the level of decentralization the current technology has. It's rather hard to have all process of a system decentralized. There is always a few people or a corporation that needs to have the final word in the developing process.

There are many barriers, some of the biggest are technological and governance, that will have to overcome. But it still gives us incredible opportunities to improve on. The most significant challenges of all will be how the blockchain will scale to be able to process the amount of transaction to be widely adopted. This will probably require a different consensus algorithm then the Proof of Work. Blockchains are still work in progress, and there is a long way for mass adoption. This is probably a good thing because no chain could scale to that level of adoption yet.

It's important to understand that it's not only possible to use it as a currency. It's not unusual to think blockchain is Bitcoin; it's more the other way around. But given the benefits of blockchain, for example, transparency and eliminations of intermediaries, the technology could give incredibly powerful area of use and could contribute to services like digital voting.

An interesting reflection is when this work began, I thought it would be tough to find any challenges and drawbacks. It's instead of the other way around now, I now see the problems more clearly and realize there is a long way until this technology could be widely adopted. One of the biggest challenges blockchain faces today is how the system will scale to be widely adopted. There is plenty of time to make changes; nothing is perfect from the beginning. Blockchains could change many industries, but it's important that people see and accept the challenges of scaling, adoption and government regulations.

Future work

For future work, it would be good to do more research and explain wallets. Do a more in-depth comparison of all the relative consensus algorithms. It would also be a good idea to find out how many people know about this technology and how many understand it on a technological level to get a grasp on how the current state actually is.

Try more dApps and get a understand of how many applications really use this technology to it’s full potential. Also to see if there is some kind of bubble, in other words people deploying dApps that really does not need to be dApps.

Reference

[1] https://bitcoin.org/bitcoin.pdf
[2] https://www.coindesk.com/data/bitcoin-market-capitalization/
[3] https://en.bitcoin.it/wiki/Transaction
[4] https://www.ibm.com/blockchain/business-use-cases.html
[5] https://www.samsungsds.com/global/en/solutions/off/nexledger/Nexledger.html
[6]https://www.forbes.com/sites/laurashin/2016/05/02/amazon-steps-up-blockchain-commitment-web-services-partners-with-digital-currency-group/
[7]
https://www.forbes.com/sites/laurashin/2016/05/02/amazon-steps-up-blockchain-commitment-web-services-partners-with-digital-currency-group/
[8] https://blockgeeks.com/guides/blockchain-applications/
[9] http://www.huffingtonpost.com/ameer-rosic-/5-blockchain-applications_b_13279010.html
[10] https://www.etla.fi/wp-content/uploads/ETLA-Raportit-Reports-74.pdf.
[11] https://blockchain.info/pools.
[12] https://eprint.iacr.org/2013/829.pdf
[13] https://github.com/ethereum/wiki/wiki/Proof-of-Stake-FAQ
[14] http://www.comp.nus.edu.sg/~prateeks/papers/Bitcoin-scaling.pdf.
[15] https://raiden.network/
[16] https://github.com/ethereum/wiki/wiki/Sharding-FAQ
[17] https://www.youtube.com/watch?v=9Z5akEKz_9Y
[18]
http://www.scmp.com/news/china/economy/article/2111456/why-has-china-declared-war-bitcoin-and-digital-currencies
[19] https://www.forbes.com/sites/quora/2017/09/21/whats-holding-blockchain-back-from-large-scale-adoption/
[20] https://hbr.org/2017/01/the-truth-about-blockchain
[21] https://github.com/ethereum/.
[22] https://github.com/bitcoin/bitcoin
[23] https://github.com/litecoin-project/litecoin
[24] https://www.huffingtonpost.com/ameer-rosic-/7-incredible-benefits-of-_1_b_13160110.html.
[25]
https://hackernoon.com/blockchains-dont-scale-not-today-at-least-but-there-s-hope-2cb43946551a
[27] https://medium.com/@VitalikButerin/the-meaning-of-decentralization-a0c92b76a274
[28] https://hbr.org/2017/04/who-controls-the-blockchain
[29] https://blockchain.info/
[30] https://www.ibm.com/blogs/cloud-computing/2017/04/characteristics-blockchain/
[31] https://getmonero.org/resources/about/

Wordlist

Node
A device that is connected to the blockchain.

Full node
A device that is connected to the blockchain that also verifies transactions. So-called mining.

dApp
A decentralized application.