Blockchain

Assumptions

1) The bitcoin is a **large** peer-to-peer network of nodes.

- Peers can join or live the network, but it's very likely they will actually stay;
- The network supports **broadcast.** This is implemented using anti-entropy. By using mechanisms to **request missing messages**, it's very likely that all nodes will receive the message.

2) Each account contains a public and private key. An account also has a id/number which is the hash of its public-key.

Blockchain

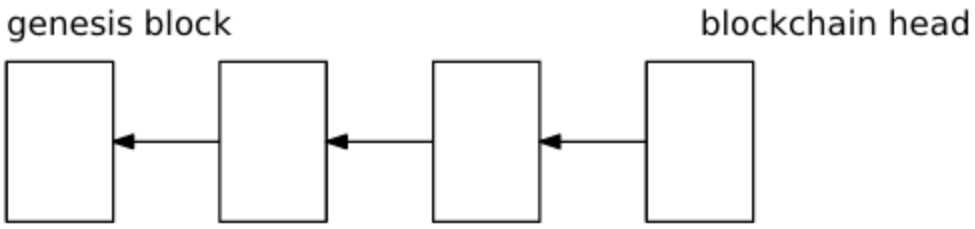
Mastering Bitcoin

The blockchain data structure is an ordered, back-linked list of blocks of transactions. The blockchain can be stored as a flat file, or in a simple database. The Bitcoin Core client stores the blockchain

https://www.oreilly.com/library/view/mastering-bitcoin/9781491902639/ch07.html

It’s a record of all transactions ever performed in a distributed fashion.

A block is a set of events, where the first block in a blockchain is called the **genesis block.** All the following blocks points to the previous block in the chain, like a linked list where the **link is just a reference to the hash** of the previous (parent) block . On this way it’s more unlikely to change the order of the blocks. The most recent node is called **blockchain header.**



Network

The peers maintains a connection to other peers.

By the specification, each nodes attempts to connect to 8 other nodes, but the degree can be much larger since there’s no limit to how many nodes one can connect to.

> Currently (09 jan 2022) the bitcoin blockchain contains 375 gigabytes.

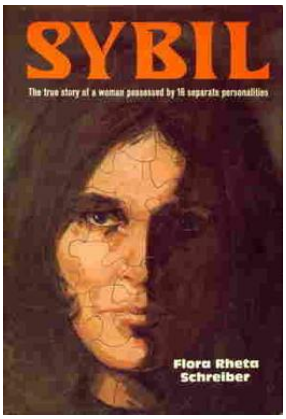
Consensus

Sybil attack - Wikipedia

A Sybil attack is a type of attack on a computer network service in which an attacker subverts the service's reputation system by creating a large number of

pseudonymous identities and uses them to gain a disproportionately large
𝕎 https://en.wikipedia.org/wiki/Sybil_attack

Name given to an attack thatcreate multiple pseudonymous identities to gain a large influence. The name was given in reference of a book called “Sybil”, where protagonist had a **dissociative identity disorder.**



Peers/Nodes need to agree if the block will be added to the chain or not and also on the order it will be added.

Since we don’t have a central authority to validate and verify the secure transactions, it’s necessary to create an alternative. In the case of the blockchain all the other nodes needs reach in a **common agreement.**

It’s a little complicated to apply the resolutions presented in the classes of byzantine failure in quorums, because it’s necessary to know the number of nodes in the network and it’s complicated to calculate in such a big network.

Proof of Work

The use of **Proof of Work (PoW)** makes the *Sybil attack* impracticable , due to a series of rules that need to be followed to create a new block. One of these rules is specified by the **Proof of Work.** Basically it makes the processing make intense use of computational resources. So, having many computers performing this action will have a great cost to the attacker.

How it works

To add a block to the chain, a computer **must solve a random cryptographic puzzle, where the solution is not trivial:** the computer must find a nonce to include in the block header, such that the header’s SHA-256 is **smaller** than a known **target** (the random number).

> Remember that SHA-256 is a NP problem, but not P. So, it must be solved by brute force.

Target

Can be tuned so as to adjust the puzzle difficulty. The expected number of hashes to solve is:

$$N_{hashes} = 2^{256}/target$$

The highest the target, the easier is to solve, since it's not difficult to find a number smaller than the a really high number.

> The bitcoin is designed to produce a block every 10 minutes. After every 2016 blocks (after 14 days more less), an adjustment is done, so that the generation time continues the same.

Miners

To generate the proof-of-work for a block, a node doesn't need to keep the entire blockchain.

The miners also needs to be able to quickly compute hash values and nowadays miners uses ASIC to mine.

> ASIC is a computer equipment with circuits and integrated chips that were developed with very specific functions.

Broadcasting

When solving a PoW, a node broadcasts the new block.

Once received a node checks its validity:

- Verify the PoW (i.e. compute the hash of the header and verifies if it's equal)
- Checks all transactions in the block

If the new block is valid:

- Stops working on the PoW
- Adds the new block to the chain
- Forwards the new block

Basically it checks a hash and if it’s valid forward the block.

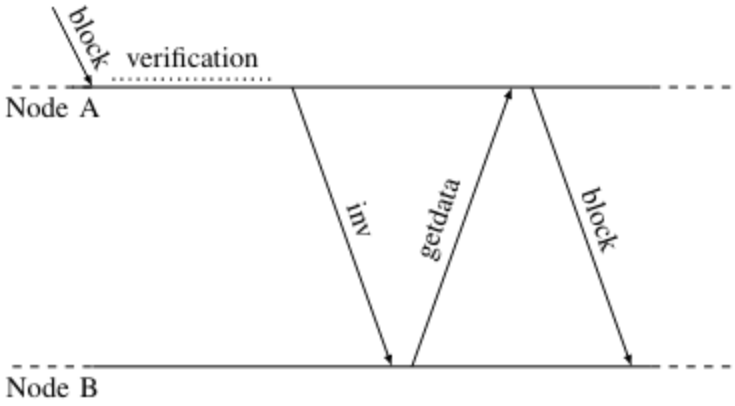
However, by the time a node receives a new block, its chain may be missing some of its parents. So, it will necessary to fetch and validate all the missing blocks. The fetch action to synchronize a node is fast, since the protocol was designed to have some efficiency.

Block broadcasting with anti-entropy

Upon the validation of a new block, a node (A) sends to its neighbors **inv(entory)** messages with a set of hashes it has.

Then other node (B) that received the **inv** message requests to (A) the blocks that it doesn’t have in its blockchain. This is the **getdata** message.

Then (A) returns to (B) the blocks asked.



Block propagation delay

This process of verification can add a significant delay to the process of mining.

The blockchain validation is repeated every hop. So for every transmission of a message this verification is performed. For this reason the validation time adds to the time to transmit a message.

Nowadays the propagation delay doesn’t take so long (less than one second). This measure is not global, other nodes (just a few), may have a delay of 10 seconds.

This improvement was caused by the improvements suggested by 2013 paper by Decker and Wattenhofer. The relevant suggestions (in order of relevance) were:

- Reduction in the diameter of the overlay network, by increasing the connectivity. By increasing the distance between two nodes the time to one message from (A) to (B), the number of validations is higher than in smaller distance, because there’re more hops.
- Faster validation, thanks to faster HW
- Minimize verification delay, by advertising a new block after the PoW verification, but before validating transactions (a lot more expensive).

Forks

Occurs when two or more nodes adds a block to their own blockchain at the same time. The fork that will persist is the one with more effort and the other is discarded.

There’s no guarantee that a block added to the chain will persist. But, the more confirmations the block has, it’s becomes more unlikely to it be dropped. Blocks with 6 confirmations are often considered final. Thus, we have here an **eventual consistency.**

But at some point Nakamoto added a “code-based checkpoint”. This is the hash of a block that cannot be replaced and thus all the blocks before it. This code-based checkpoint is hardcoded in the software.

Forks can also be generated by network partitions.

Issues

Scalability

- Blocks can be 1MB long.
- PoW is computationally intensive and increases by the system Hash-Power. The target is one block every 10 minutes.

Transaction Rate Bound

- It can do less than 8 transactions per second. But visa can do 1700 transaction per second.
- If we increase the **block size** the **block propagation** will increase and the **size of the blockchain** would increase at a rate of 500GB/year

- If we increase the block rate to 1 per minute, **forking** will be much more frequent.

Energy consumption

- To be secure the computers need to have a huge hash-power and the PoW is tuned so as to ensure a constant block-rate. **This consumes a lot of energy.**
- If it consumes a lot of energy, then we may also be generating impacts in the **climate change** and there is an **opportunity cost.** The energy expended mining could be used for another purposes.

Proof-of-Stake

(more than we need to know, but the slides weren't that complete....)

Proof of stake - Wikipedia

Proof of stake (PoS) protocols are a class of consensus mechanisms for blockchains that work by selecting validators in proportion to their quantity of holdings in the associated

https://en.wikipedia.org/wiki/Proof_of_stake

IEEE Xplore Full-Text PDF:

https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8746079

It's an alternative to PoW. Ethereum plans to replace PoW with PoS and it's already in use by Cardano.

Here participants with higher coinage i.e the integral of the number of coins by their holding time, have higher chance of being selected to create the next block and add it to the blockchain. Now the block leaders are selected not by their computational power but by their coin age.

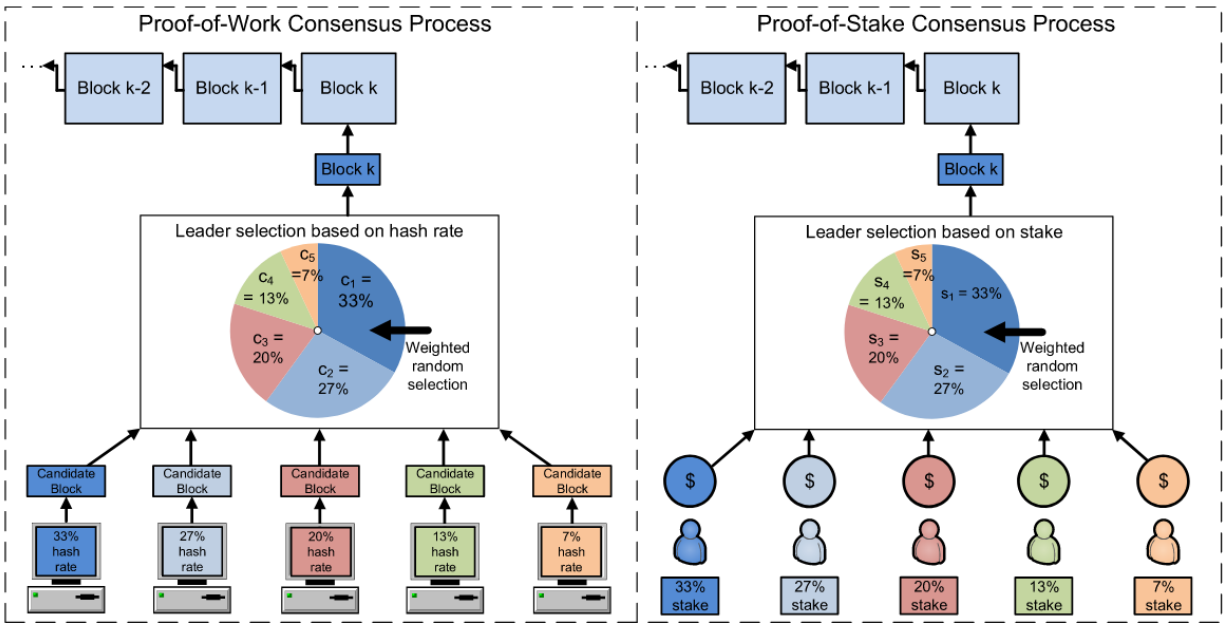
Thus, it's not necessary to spend a lot of energy searching for leaders like in PoW. Using index, the FTS algorithm searches the transaction history to find and select the current owner of that token to be a leader. So, the probability of a participant being

chosen to add the new block is represented by the equation on the right.

$$p_i = \frac{s_i}{\sum_{j=1}^{N} s_j},$$

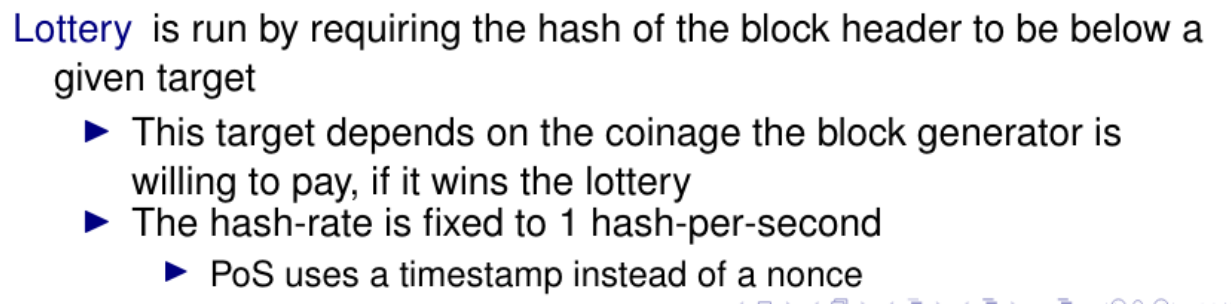
N - number of participants; s - stakes (can be coinage).

This image represents the probability:



However, some clock synchronization is needed to validate the blocks. Given the propagation delay, using NTP is more than enough.

When a block is recently added to the blockchain, the owner coinage is not update right away, because the blockchain generated may not be final.



The disadvantage of **PoS** appears to be:

- It is harder to get right: The replacement of PoW by PoS in Ethereum was delayed many times.
- Have some undesirable properties to implement a decentralized crypto-currencies: check PoX.

Smart contracts

> Smart contract - Wikipedia
>
> A smart contract is a computer program or a transaction protocol which is intended to automatically execute, control or document legally relevant events and actions according to the terms of a contract or an agreement. The objectives of smart contracts are the reduction of need in trusted intermediators, arbitrations and
>
> W https://en.wikipedia.org/wiki/Smart_contract

- It allows to store unforgettable data in persistent and transparent way.
- It can be used to implement **smart contracts.**

Are programs stored on a blockchain that run when predetermined conditions are met. They are used to reduce the number of trusted intermediates and other things.

PBFT

PBFT and Bitcoin

PBFT can be used to maintain a replicated log, i.e. a blockchain, but

- It was designed for 4 or 7 replicas, i.e. for $f = 1$ or $f = 2$
- Its complexity in terms of messages is $O(n^2)$

Bitcoin showed that PoW scales (kind of) to thousands of nodes

- But this may not be needed in permissioned networks

Taxonomy regarding authorization to maintain, grow and access a **blockchain**

Permissionless/Public anyone can both read and grow the blockchain. Example: Bitcoin

Consortium maintained by the members of a consortium

- Each member may run a few nodes
- These nodes are responsible for persistence but also for growing the blockchain
- The blockchain may use different read policies, i.e. from open access to consortium-only

Private/Permissioned owned by a single organization, which controls which blocks are added