Blockchain

Bitcoin

1. There is a (large) peer-to-peer network of nodes with some computing resources
2. There is a set of accounts each of which has a pair of private and public keys

Bitcoin Blockchain

Blockchain: sequence of blocks
Block: contains a set of transactions

- Each block contains a **header** with metadata
- The first block in the chain is the **genesis block**
- Blocks are appended to the **blockchain head**

Network

- Bitcoin's blockchain is maintained by a peer-to-peer network
- Peers maintain random connections to other nodes/peers
- Peers maintain a copy of the entire blockchain

Consensus

- Consensus is needed to agree on the blocks and on their order
- Conventional Byzantine Algorithms either Byzantine Quorums or PBFT rely on quorums, i.e. sets of nodes, but in a P2P network
    - It is difficult to know how many nodes there are
    - Worse, it is fairly easy to create multiple identities

Bitcoin Proof-of-Work (PoW)

Solve a cryptographic puzzle that takes a random but large time

- SHA-256 is a non-invertible function, thus this puzzle must be solved by brute force
- Target can be tuned so as to adjust the difficulty of solving the puzzle

Miners

The block header includes, among other metadata:

- The hash of the previous block

- The hash of the remaining of the block i.e. the transactions

To generate the proof-of-work for a block, a node needs not to keep the entire blockchain
Thus the PoW is computed by nodes, miners, which nowadays use ASIC's specially designed for Bitcoin

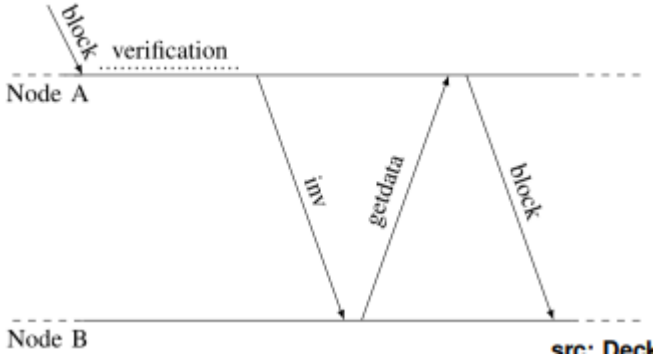
Block Broadcasting

- Upon solving the PoW, a node broadcasts the new block
- Upon receiving a new block, a node:
    - Checks its validity
        - Verifying its PoW, i.e. computing the hash of its header
        - Checking all transactions in the block
    - If the node is valid:
        - The node stops working on the PoW for a block extending the current head
        - Adds the new block at the head of the blocchain
        - Forwards the new block
- In both cases, a node starts working on the next block, which will follow the one just added
- When a node receives a new block, its chain may be missing some of its ancestors
    - The node will have to fetch and validate the missing blocks
    - The protocol is designed to efficiently synchronize nodes that were disconnected for some time

Block Broadcasting with Anti-Entropy

- **Upon validation of a new block** a node sends to its neighbors **inv(entory)** messages with a set of hashes of blocks it has
- **Upon receiving an inv message** with hashes of blocks it does not have in its blockchain, a node sends a **getdata** message with a list of the hashes of blocks it wants
- **Upon receiving a getdata message** a node sends each block in getdata's **block** list in its own block message
- **Each block is inserted into the network** by a miner using an unsolicited block message to one or more peers

- The block has just been generated



Block Propagation Delay

- Block validation can add a significant delay
- Block validation is repeated at every hop
- Block propagation delay has a long tail distribution

Bitcoin Forks

Fork: occurs when 2 or more nodes add a different block at the head of an otherwise identical blockchain at more or less the same time

- Resolution is based on the expected amount of work (usually the length) of competing blockchains
- Eventual consistency with high probability

Analysis

- Accidental forks depend on:
    1. The expected time to generate the PoW
    2. The block propagation delay
- selfish mining strategies may exacerbate the problem
- Network partitions can also lead to forks

Bitcoin Scalability and Energy Consumption

Scalability Issues

- Proof-of-work is computationally intensive
- Blocks cannot be larger than 1 MB long
- Storage of the whole blockchain kept by all (full-)nodes

Transaction Rate Bound

- **Bitcoin parameter tunning** cannot make for this difference of more than 3 orders of magnitude (assuming capacity of 10K)

- **Block size** if we increase it by an order of magnitude
  - **Block propagation** will increase, but may be this is OK, as we would get back to the numbers of 10 years ago
  - But block chain size will increase at a rate of 500 GB/year
- **PoW difficulty** if we increase block rate to 1 per minute (one order)
  - Forking will be much more frequent
  - This is made worse if we try to tune both block size and rate

Energy Consumption

- Extremely low energy-efficiency
- Affects Climate Change

Proof-of-Stake (PoS)

- alternative to PoW
- **Idea**: run a lottery to decide which user adds the next block to the chain
- **Coinage** (from "coin" + "age") is the product of the amount of coins by the time that amount is held
- **Lottery** is run by requiring the hash of the block header to be below a given target
- **Clock synchronization** is needed to validate blocks
- **Ties** are broken using the block's coinage
- **Coinage consumption** occurs when a block is added to the chain

Advantages

PoS is more energy efficient and has higher block-rate

Disadvantages

PoS appears to:

- Be harder to get right: The replacement of PoW by PoS in Ethereum has been postponed several times
- Have some undesirable properties to implement a decentralized crypto-currencies: check alternative PoX

Permissioned Blockchains

- Smart contract is just code that may be executed upon some event added to the blockchain
- Some applications need not to be open to the Internet at large
- Problem is mainly of ensure consensus on the contents of each block an on their order

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