*CS*06201*a*01: Network Computing and Efficient Algorithms

Lecture 11: Blockchain

Xiang-Yang Li and Xiaohua Xu

School of Computer Science and Technology
University of Science and Technology of China (USTC)

August 31, 2021

Blockchain Defined

Simply defined a Blockchain is little more than a:

- Distributed
- Secure
- Logfile

What is Bitcoin

- A protocol that supports a decentralized, pseudo-anonymous, peer-to-peer digital currency
- A publicly disclosed linked ledger of transactions stored in a blockchain
- A reward driven system for achieving consensus (mining) based on Proofs of Work for helping to secure the network
- A scare token economy with an eventual cap of about 21M bitcoins

Features of Bitcoin

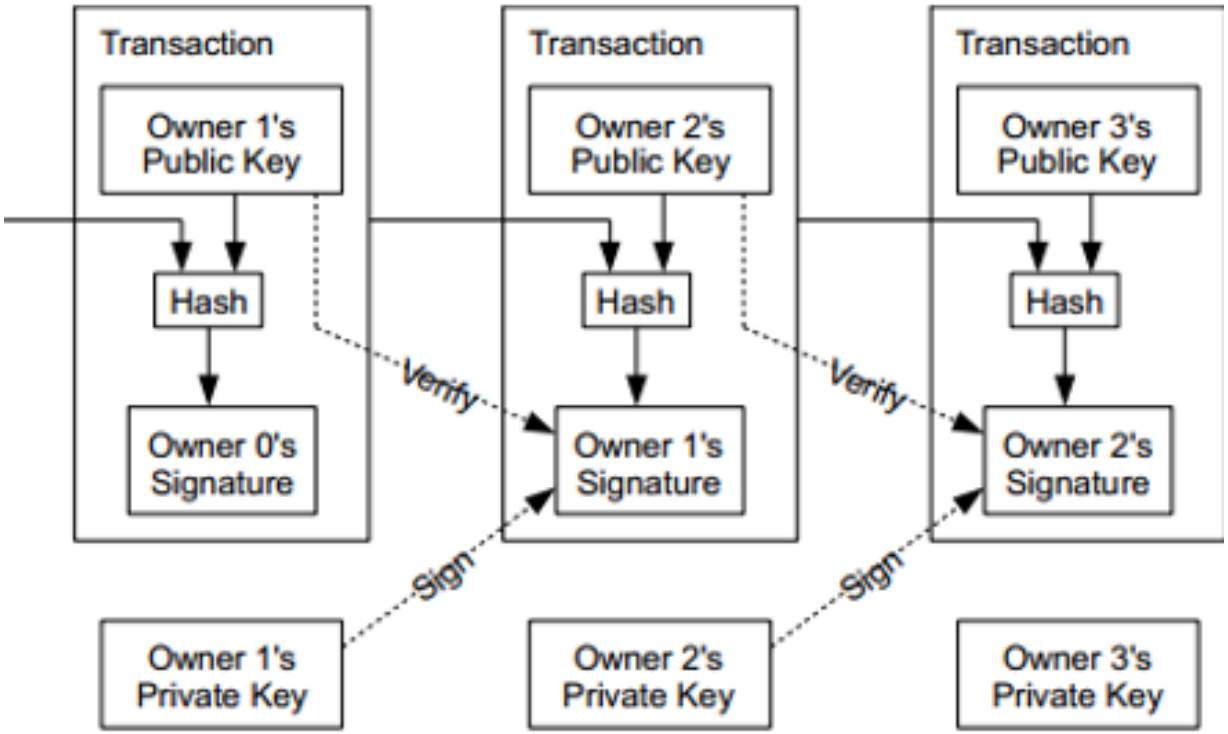
- Essentially its deflationary the reward is cut in half every four years, and tokens can be irrevocably destroyed
- Nearly infinitely divisible currency units supporting eight decimal places 0.00000001 (known as a Satoshi or Noncent*)
- Nominal transaction fees paid to the network Same cost to send $.01 as 1{,}000{,}000$
- Consensus driven no central authority
- Counterfeit resilient
  - Cannot add coins arbitrarily
  - Cannot be double-spent
- Non-repudiation aka gone baby gone no recourse and no one to appeal to return sent tokens

Decentralized

- The digital wallet operates in a peer to peer mode
- When it starts it bootstraps to find other wallets
  - Originally it used the Internet Relay Chat (IRC) network
  - Now based on DNS and seed nodes
- The wallet will synchronize with the network by downloading ALL of the transactions starting from the GENESIS block if necessary
  - 338,540 blocks at time of slide prep
  - Just over 20 GB
- Using a gossip protocol the wallets share all transaction information with their peers

Coins flow from Inputs to Outputs

A coin owner transfers coins by digitally signing (via ECDSA) a hash digest of the previous transaction and the public key of the next owner. This signature is then appended to the end of the coin.

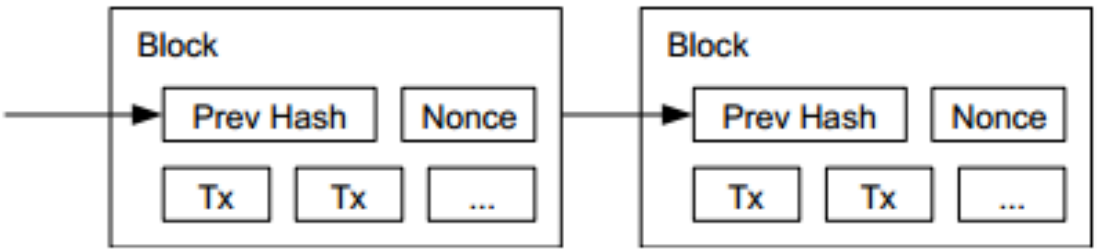


Addresses are like Accounts

- The wallet listens for transactions addressed to any of its public keys and in theory is the only node that is able to decrypt and accept the transfer
- Coins are sent by broadcasting the transaction to the network which are verified to be viable and then added to a block
- Keys can represent a MULTI-SIG address that requires a N of M private keys in order to decrypt the message

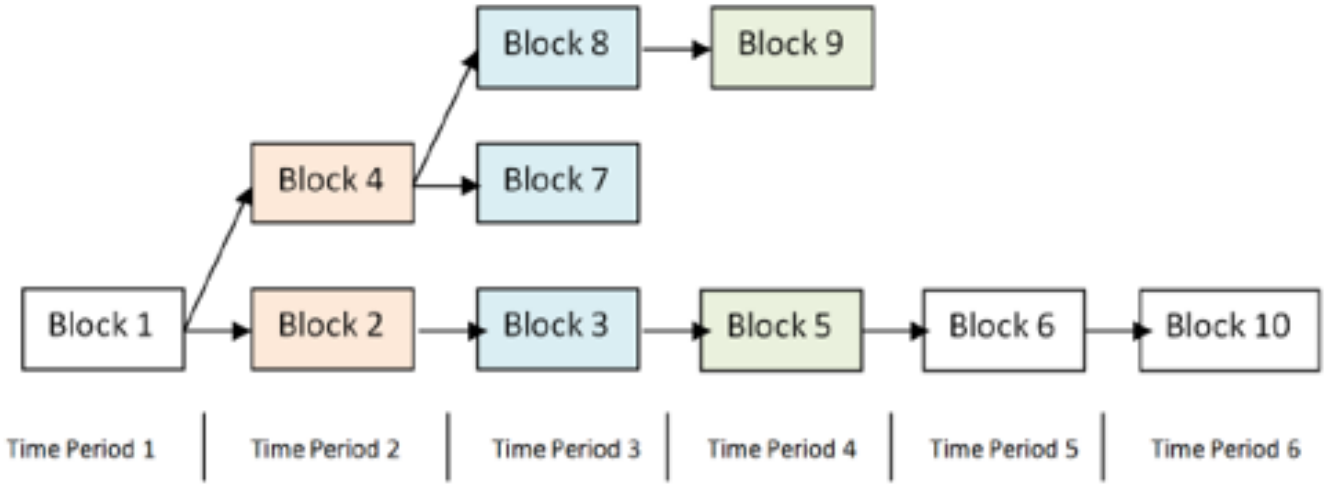
Public Ledger

- Every viable transaction is stored in a public ledger
- Transactions are placed in blocks, which are linked by SHA256 hashes.
- `https://blockchain.info`



Arriving at Consensus

- Although the accepted chain can be considered a list, the block chain is best represented with a tree.
- The longest path represents the accepted chain.
- A participant choosing to extend an existing path in the block chain indicates a vote towards consensus on that path. The longer the path, the more computation was expended building it.



Consensus Process = Mining

- Originally the digital wallet could also participate in the consensus process by attempting to secure the network directly
- This process is known as mining
- Mining involves attempting to find a numerical value, known as a nonce that when combined with all open transactions can be hashed into a value that satisfies a certain difficulty
- Custom, purpose built-hardware has long since replaced the function such that its no longer productive for simple CPU based systems to compete in the mining process, and thus it was removed

Hashcash (Or How to Pay a Byzantine Generals Salary)

- Like many great ideas to become realized, it takes a confluence of other great ideas
- Based on the idea of HashCash, a Proof of Work concept invented by Adam Back in 1997 (`http://www.hashcash.org/papers/hashcash.pdf`) Originally proposed as an anti-spam throttling mechanism
- The core idea is that before accepting a transaction, the sender must first demonstrate a cost via a computationally hard problem that can simultaneously be easily verified. This generally referred to as a Proof of Work

$$\begin{cases} \mathcal{C} \leftarrow \mathsf{CHAL}(s, w) & \text{server challenge function} \\ \mathcal{T} \leftarrow \mathsf{MINT}(\mathcal{C}) & \text{mint token based on challenge} \\ \mathcal{V} \leftarrow \mathsf{VALUE}(\mathcal{T}) & \text{token evaluation function} \end{cases}$$

The Role of Hashing

- A hash function is any function that can be used to map digital data of arbitrary size to digital data of fixed size, with slight differences in input data producing very big differences in output data.
- MD5, SHA1, SHA256
- For example, the MD5 hashes of abc compared to abC

abc
0bee89b07a248e27c83fc3d5951213c1
abC
2217c53a2f88ebadd9b3c1a79cde2638
The Quick Brown Fox Jumped Over the Lazy Dog
2dfd75162490ed3b4c893141f9ab37cf

Proof of Work

- A publicly auditable cost-function can be efficiently verified by any third party without access to any trapdoor or secret information.
- A fixed cost cost-function takes a fixed amount of resources to compute. The fastest algorithm to mint a fixed cost token is a deterministic algorithm.
- A probabilistic cost cost-function is one where the cost to the client of minting a token has a predictable expected time, but a random actual time as the client can most efficiently compute the cost-function by starting at a random start value. Sometimes the client will get lucky and start close to the solution.

The Hash Lottery

- Hashing is straightforward, but not challenging
- Unless the goal is to say, find me a hash value that satisfies a certain level of difficulty
- For example, lets say the challenge is find a hash-value that begins with a number of zeros, for a given input
- The Proof of Work comes from finding a number (known as a NONCE) that when added to the input changes the output of the hash value to satisfy the difficulty.
- In the Bitcoin world this is what mining is and in effect is little more than a lot of hash-power spent on guessing winning lottery numbers that satisfy the difficulty of the problem in order to obtain the reward from the network

The Payout

- The node that finds the best solution to the challenge is provisionally granted a reward
- Originally in Bitcoin it was 50 new coins
- Competing solutions are evaluated based on which node offers the higher number of transactions included in the candidate block as well as the level of over-satisfying the difficulty.
- For example, if two nodes offer a solution to the challenge and both have the same number of transactions, the reward will go to the node that found a NONCE that beat the challenge
  - E.G. Find a hash that begins with 4 zeros
  - The node that supplies a hash that has 5 zeros beats the node that only finds the minimum

Transaction Confirmation

- Having a transaction provisionally accepted into a candidate block signals that the network has verified that the inputs were viable
- Every new block accepted into the chain after the transaction was accepted is considered a confirmation
- Coins are not considered mature until there have been 6 confirmations (basically an hour assuming a 10 minute block cadence)
- New Coins created by the mining process are not valid until about 120 confirmations
- This is to assure that a node with more than 51% of the total hash-power does not pull off fraudulent transactions

Why 51% Matters

- When does 1 + 1 = 3 ? *
- In the case of Bitcoin consensus goes to the chain with the highest number of blocks
- Not just in theory, but in practice several large mining pools have generated six blocks in a row
- To date the network has voluntarily shifted its mining power around or faced Distributed Denial of Service attacks
- When everyone says it does!