BLOCKCHAIN AND DISTRIBUTED LEDGER TECHNOLOGIES

Lecture 2 – Introduction to the Bitcoin Blockchain

Youngwhan „Nick“ Lee
DeLight Chain, Inc.

WORLD BANK GROUP

VEERMATA JIJABAI TECHNOLOGICAL INSTITUTE
V.J.T.I.
ESTD 1887
MUMBAI 400019

DeLight Chain

HAW HAMBURG

LEARNING OUTCOUMES

After this lecture students shall

- **know**
  - what Bitcoin is
  - the historical importance and development
  - about the protocol steps and special roles in Bitcoin
  - the concepts of double spending detection and avoidance
  - mining and consensus

- **can describe**
  - a full way of a Bitcoin transaction
  - evolution steps and major types of BC/DLT systems

- **Can set up** a full Bitcoin node and operate a transaction

BITCOIN INTRODUCTION

*"A purely **peer-to-peer** version **of electronic cash** would allow **online payments** to be sent **directly from one party** to another **without going through a financial institution**. **Digital signatures** provide part of the solution, but the main benefits are lost if a trusted third party is still required to **prevent double-spending**.*
[…]
***The network timestamps transactions by hashing** them into an ongoing chain of hash-based **proof-of-work**, forming a record that **cannot be changed** without redoing the proof-of-work.*

*The **longest chain** not only serves as proof of the sequence of events witnessed, but **proof that it came from the largest pool** of CPU power. As **long as a majority of CPU power is controlled** by nodes that are not cooperating to attack the network, they'll generate the longest chain and outpace attackers. […]"*

Source: Satoshi Nakamoto: Bitcoin: A Peer-to-Peer Electronic Cash System, 2008

BASIC CONCEPTS

- **identity**
  the Bitcoin address for every user

- **privacy**
  anonymity of the identities in the Bitcoin network is given that only transactions are made public

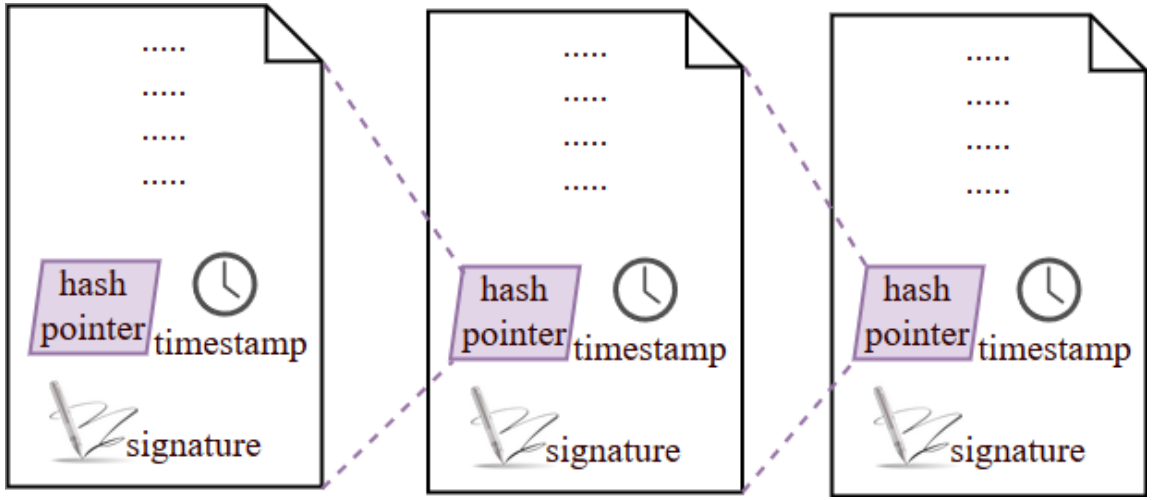
- **block building and consensus**
  the process of validate new blocks and adding them to the Bitcoin blockchain by using individual CPU time and electricity (to solve mathematical puzzles)

Record transactions in a ledger

- timestamp a document (= bring into an order with other documents) and to issue the „certificate“ of it
- goal:
  give the time of when a document came into existence

By doing so, to accurately convey the order of creation of documents

- required: a document‘s timestamp shouldn‘t be changed after the fact
- The certificate of a document ensures the integrity of the contents of the previous document



A document‘s certificate signs the three data elements

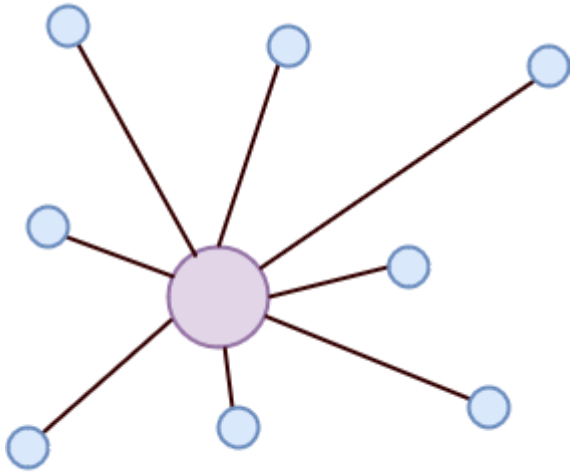
- hash pointer to previous document‘s certificate,
- current timestamp and
- document data together.

Sources/ Further reading:

Haber, S., Stornetta, W.: Secure names for bitstrings. CCS, (1997)

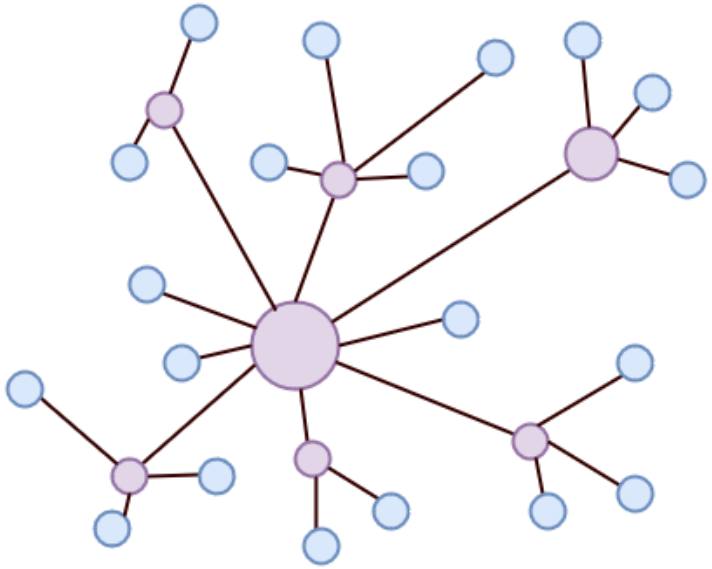
TYPES OF NETWORK

Three Kinds of Network



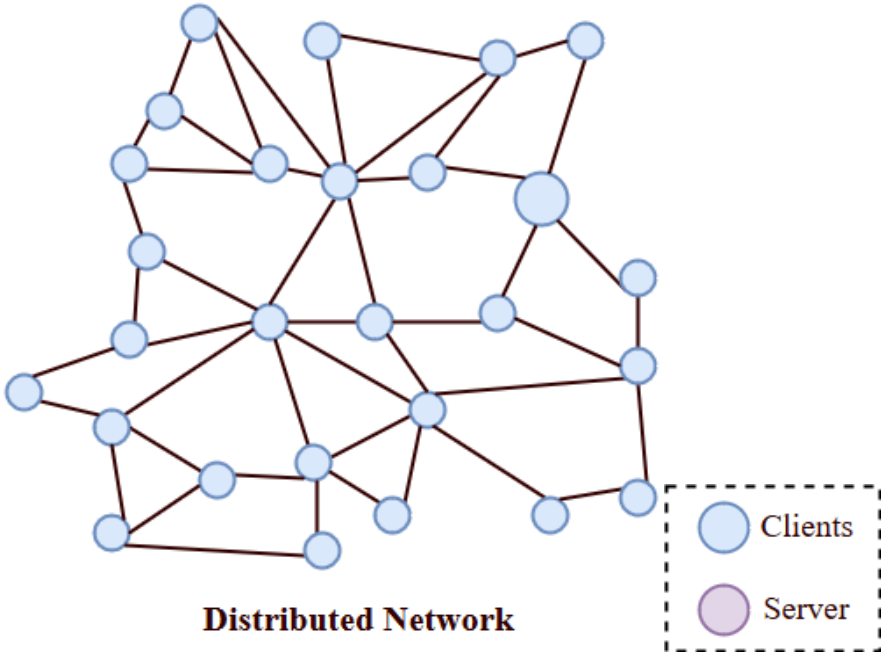
Centralized Network

In a centralized network, individual clients request services and recources from centralized servers.



Decentralized Network

In a decentralized network, local servers offer services and resources for individual clients nearby.



Distributed Network

In a distributed (or P2P) network, interconnected nodes („peers“) share resources each other without the use of a centralized administrative system.

DECENTRALIZATION IN BLOCKCHAIN

- **Logical (de)centralization**: Do the systems interface and data structure look like a single huge object or like an amorpous swarm? If both, providers and users of the system, cut in half, will both halfes be able to fully operate as independet units?
- **Political (de)centralization**: By how many individuals or organizations, that ultimately controll the computers, is the system made up of?
- **Architectural (de)centralization:** If the system is breaking down, how many physical computers can it tolerate at any single time? And of how many physical computers is the system made up at all?

➡ Blockchains should be

- Politically decentralized – no one controls them
- Architectural decentralized – no infrastructural central point of failure
- Logically centralized – one commonly agreed state; system behaving like a single computer

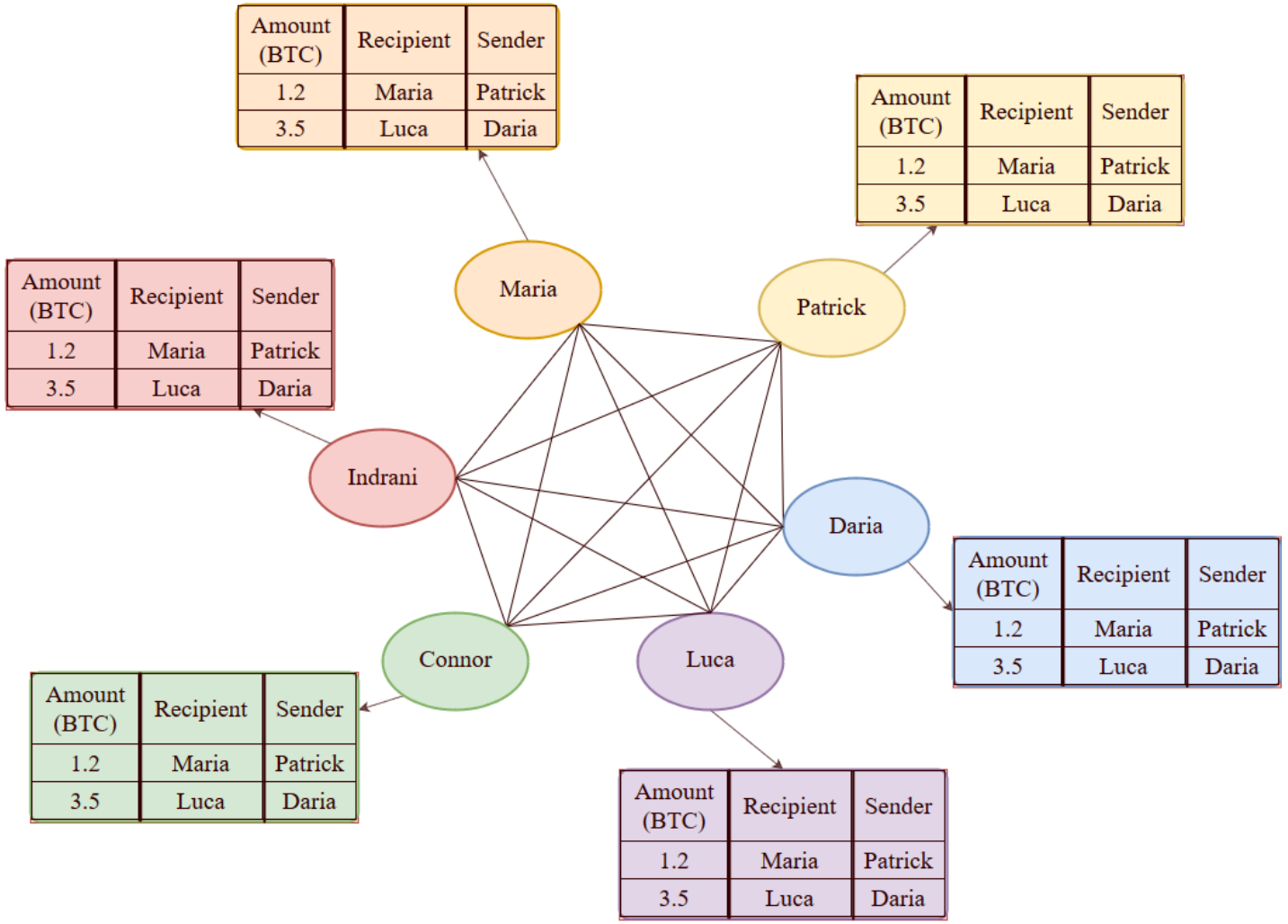
The system must prevent participants from any kind of **cheating**

- The assumption: it is a cyberspace that uncertain number of anonymous people do economic activities

- **How** can we **store all the information** and keep the integrity of the ledger of transactions?

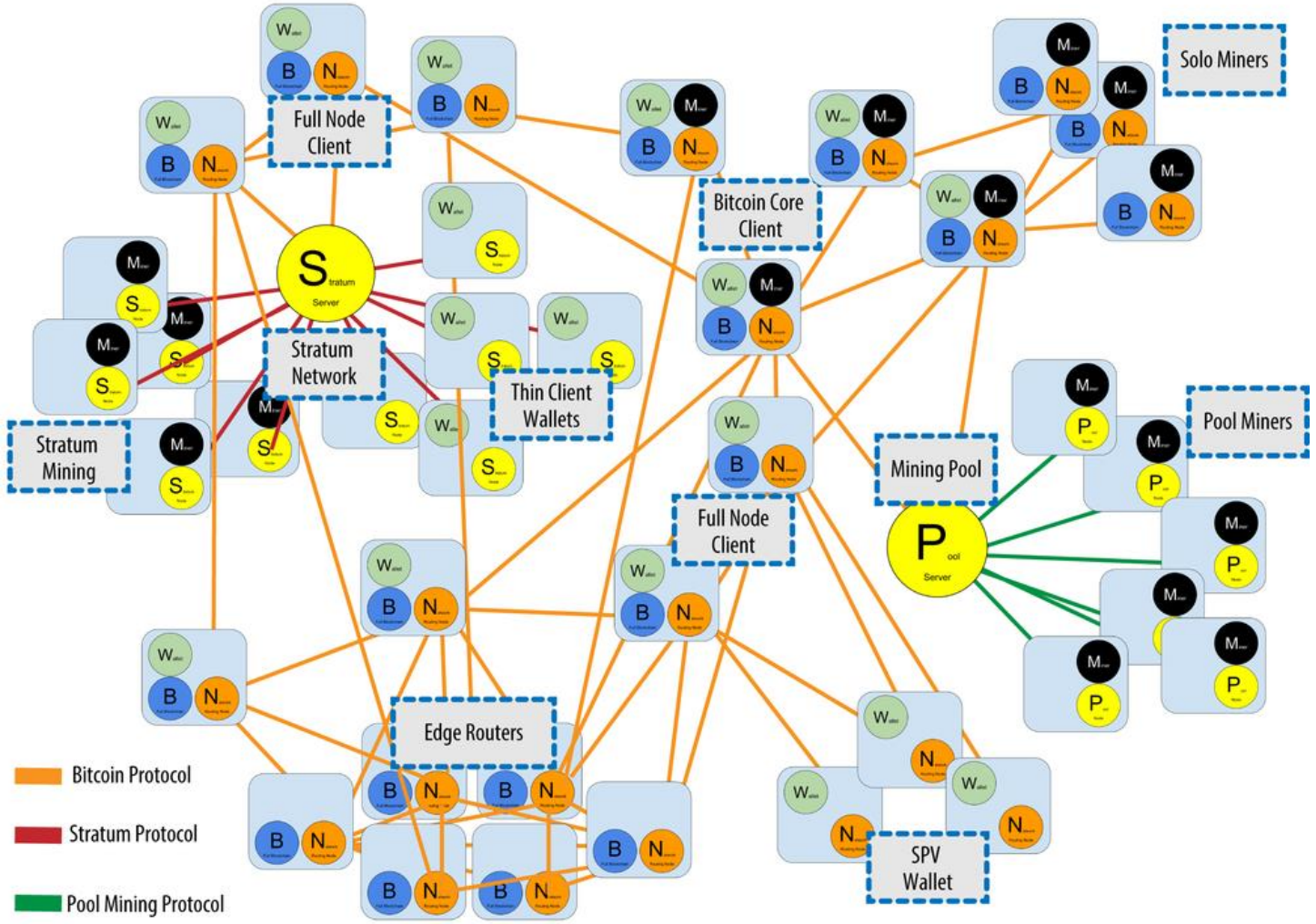
➡ Simple but smart answer: Let **everyone keep the ledger**.
Hard part: Make sure everyone has the exactly the same ledger at any given time.

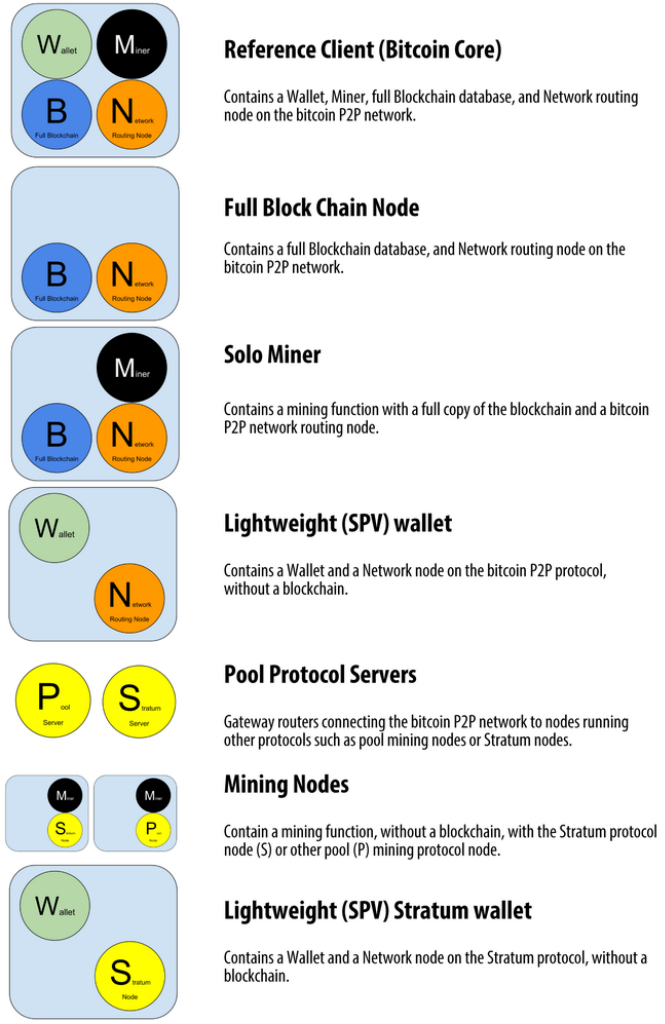
DISTRIBUTED LEDGERS



➡ Everyone stores the ledger!

THE BITCOIN NETWORK ROLES OVERVIEW

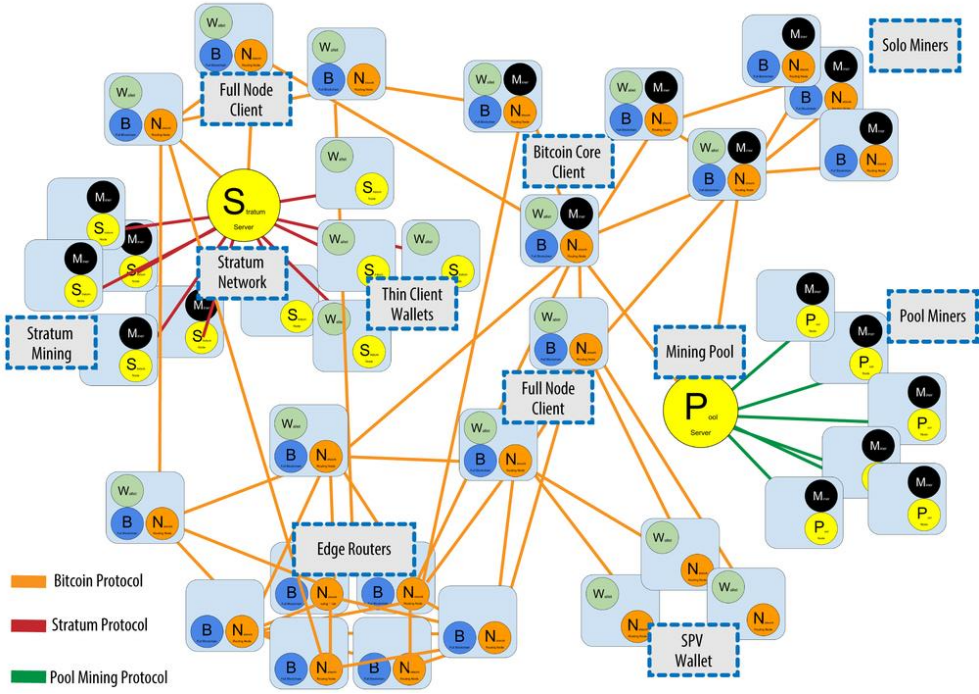




source: [Mastering Bitcoin by Andreas M. Antonopoulos](), O'Reilly (2015)

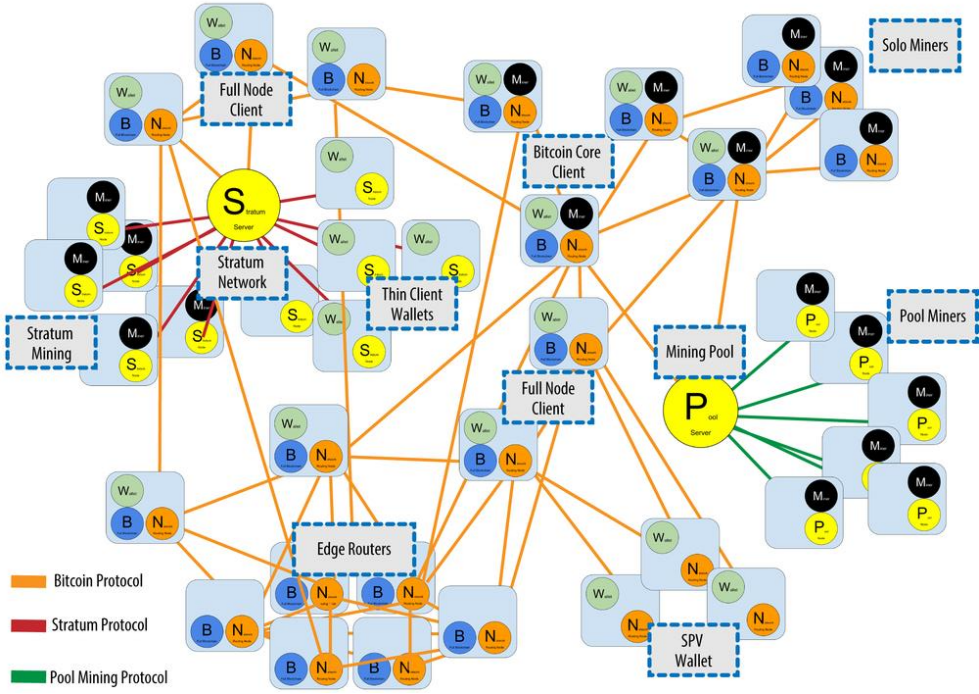
THE BITCOIN NETWORK ROLES NODES

- **Nodes**
  - any computer (or enddevice) that connects to the Bitcoin network with a collection of services: routing, blockchain database, mining, and wallet
  - validate, propagate transactions,
  - discover and maintain connections to peers
  - **Full Nodes:** storing the complete, up-to-date Bitcoin blockchain
  - **Light Nodes:** storing only parts of the Bitcoin blockchain and perform simplified payment verification (SPV) only



THE BITCOIN NETWORK ROLES WALLET

- **Wallets**
  - a collection of private keys of an Bitcoin network address to manage those keys and make transactions and store transactions from and to this address
  - Light wallets do not carry a copy of the blockchain but perform simple payment verifications (SPV):
    - only downloads block-header, 1000 times smaller than transactions
    - search for block containing transactions and looking for the single transaction in Merkle tree
    - if block is older than 6 blocktimes, it counts as confirmed



THE BITCOIN NETWORK ROLES MINERS

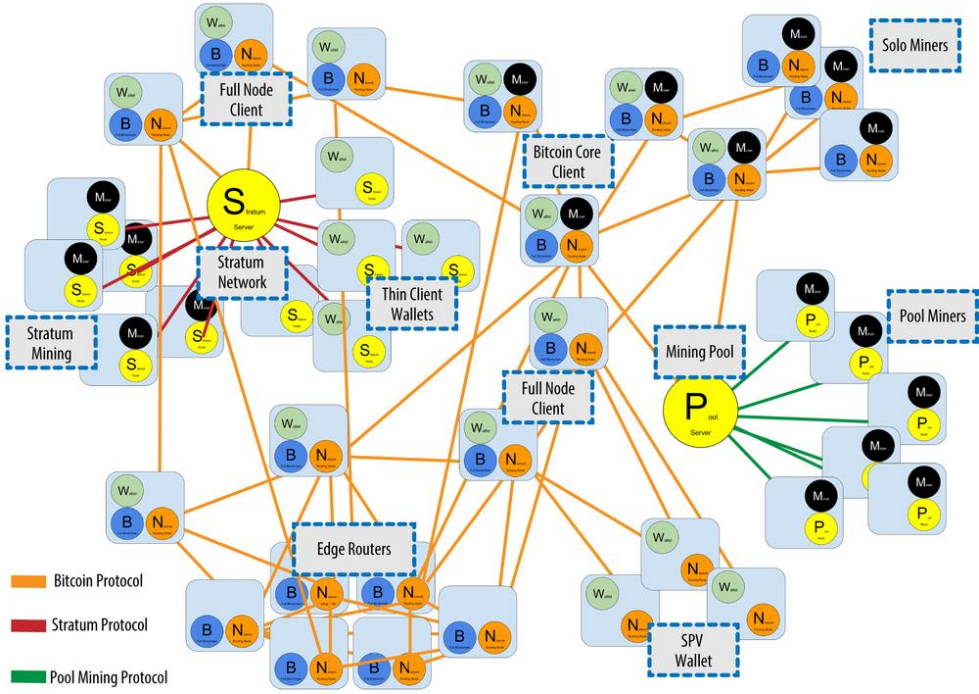
- **Miners**
  - full node that uses CPU time to solve the proof-of-work-puzzle
  - Create blocks using collected valid transactions and solve puzzles as „**Proof-of-Work**“ mechanism
  - Verify incoming transactions by checking signatures and confirming the existance of the valid bitcoins
  - Find a valid nonce (number used once) to create a valid block header with a defined number of leading zeroes (= difficulty) satisfying the following equation

$$H(nonce||prev_{hash}||transaction||transaction||\ ...\ ||transaction) < target$$

  - Equation for difficulty:

$$difficulty = difficulty * \frac{two_{weeks}}{time\ to\ mine\ prev\ 2016\ blocks}$$

→ the faster miners become, the less time is needed per block, the higher the difficulty



INCENTIVES FOR MINING

- PoW is computation intensive and consumes a lot of energy
- for motivation, the miners achieves incentive = Block reward + transaction fees

  Block reward currently is 12.5 BTC.

  The block reward halfes every 210.000 blocks.

→ amount of total Bitcoin emitted is limited to 21.000.000

→ mining is the only way to (temporarily still) increase number of Bitcoins

Reward-Drop ETA date: **30 May 2020 09:16:12**

The Bitcoin block mining reward halves every 210,000 blocks, the coin reward will decrease from 12.5 to 6.25 coins.

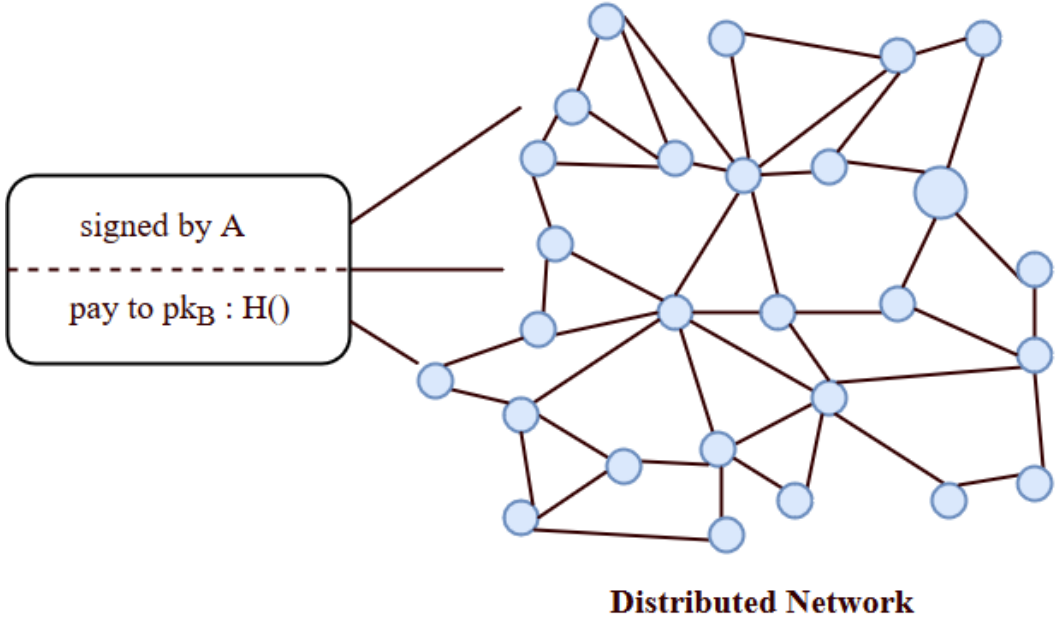
Total Bitcoins in circulation:	16,971,963
Total Bitcoins to ever be produced:	21,000,000
Percentage of total Bitcoins mined:	80.82%
Total Bitcoins left to mine:	4,028,038
Total Bitcoins left to mine until next blockhalf:	1,403,038
Bitcoin price (USD):	$6,904.00
Market capitalization (USD):	$117,174,429,100.00
Bitcoins generated per day:	1,800
Bitcoin inflation rate per annum:	3.95%

http://www.bitcoinblockhalf.com/

DISTRIBUTED CONSENSUS PROTOCOL IN BITCOIN (1/2)

Definition: There are n nodes that each have an input value. Some of these nodes are faulty or malicious. A **distributed consensus protocol** has the following two properties:

- It must terminate with all honest nodes in agreement on the value
- The value must have been generated by an honest node



Broadcasting a transaction:
In order to pay B, A broadcasts the transaction to the entire Bitcoin peer-to-peer network.

DISTRIBUTED CONSENSUS PROTOCOL IN BITCOIN (2/2)

The nodes in the system must agree on exactly

- which transactions were on broadcast and
- the order in which these transactions happened.

→ This results a single, global ledger for the system.

Problems:

- Imperfections in the network, such as latency and node crashing
- Deliberate attempts by some nodes to disturb the process

DECENTRALIZATION IN BLOCKCHAIN (REPEATED)

- **Logical (de)centralization**: Do the systems interface and data structure look like a single huge object or like an amorpous swarm? If both, providers and users of the system, cut in half, will both halfes be able to fully operate as independet units?
- **Political (de)centralization**: By how many individuals or organizations, that ultimately controll the computers, is the system made up of?
- **Architectural (de)centralization:** If the system is breaking down, how many physical computers can it tolerate at any single time? And of how many physical computers is the system made up at all?

➡ Blockchains should be

- Politically decentralized – no one controls them
- Architectural decentralized – no infrastructural central point of failure
- Logically centralized – one commonly agreed state; system behaving like a single computer

THE BITCOIN NETWORK ROLES NODES

GLOBAL BITCOIN NODES DISTRIBUTION

Reachable nodes as of Wed Apr 11 2018 22:48:35 GMT+0200.

11033 NODES

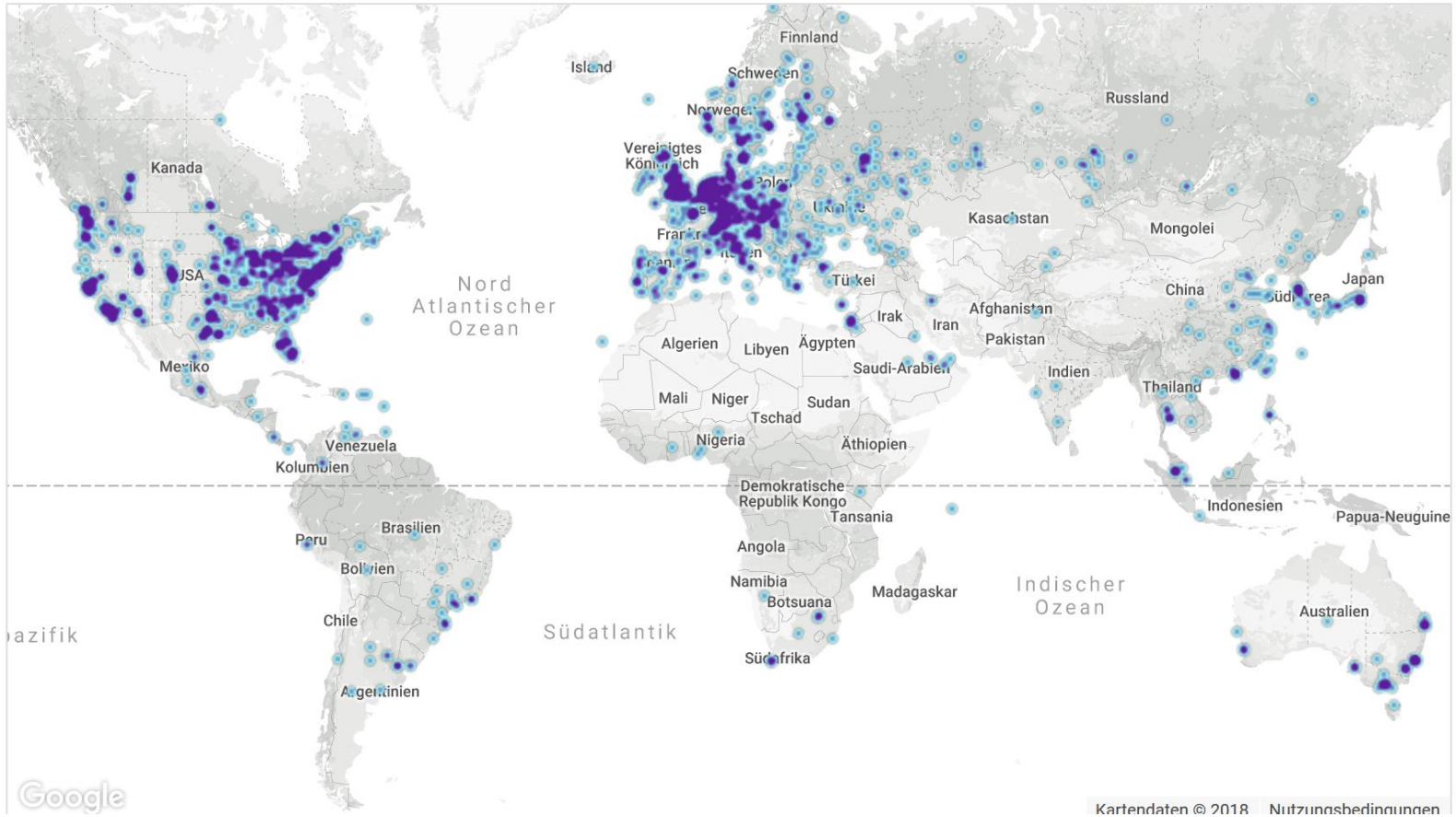
24-hour charts »

Top 10 countries with their respective number of reachable nodes are as follow.

RANK	COUNTRY	NODES
1	United States	2637 (23.90%)
2	Germany	2017 (18.28%)
3	China	1076 (9.75%)
4	France	668 (6.05%)
5	Netherlands	499 (4.52%)
6	United Kingdom	398 (3.61%)
7	Canada	390 (3.53%)
8	Russian Federation	360 (3.26%)
9	n/a	324 (2.94%)
10	Singapore	236 (2.14%)

More (104) »

https://bitnodes.earn.com/



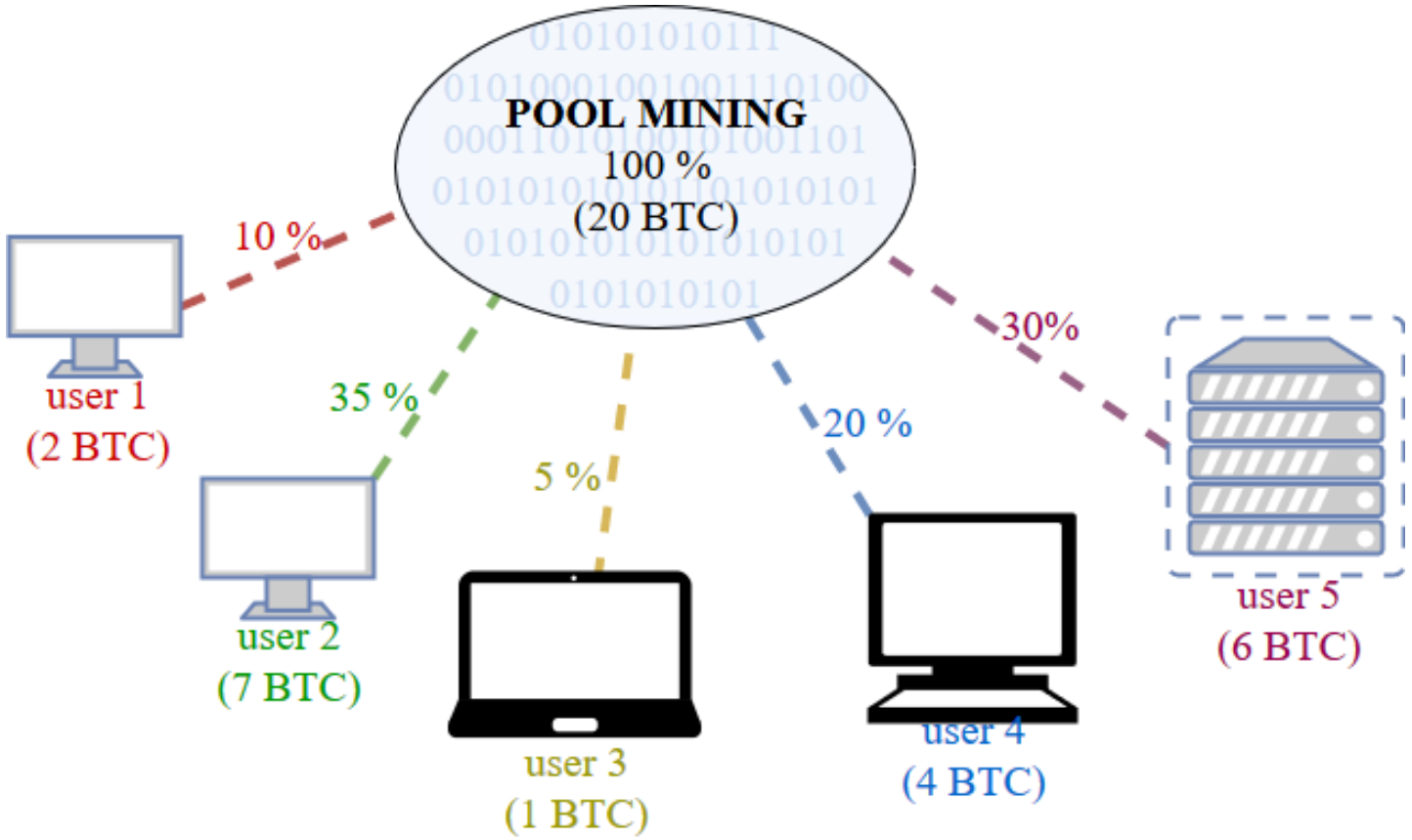
MINING POOLS

Pros

- Allows individual miners to participate
- Easy to upgrade software changes

Cons

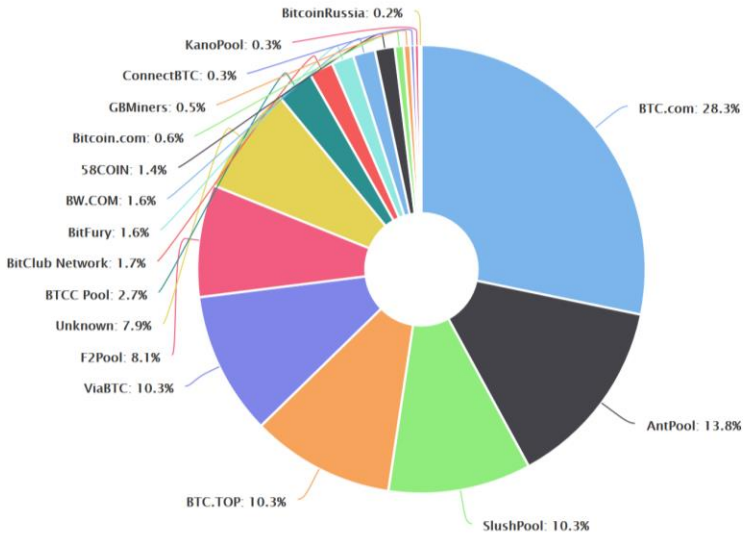
- Pool manager must be trusted
- Centralized
- Enables a multitude of attacks



THE BITCOIN NETWORK ROLES MINERS

https://investoon.com/charts/mining-map/btc, 11.4.2018

https://blockchain.info/de/pools
11.4.2018



REAL WORLD MINING (4/4)

Quick facts

- Today's network hashrate: 7,257,882 TH/s
- Mining Reward / yr = (1 yr / 10 mins) * 12.5 = 657k BTC / yr
- Assume constant price of $4000

Suppose you want to start mining today.

- Antminer S9: Costs $3000, 14 TH/s
- % of network hashrate = (14 TH/s) / (7,257,882 TH/s) = 0.000192893%

Expected Annual Reward

- 0.000192893% * 657k /yr ≈ 1.27 BTC / yr ≈ $5080/yr

Solo mining

- 1 block mined per 571853 blocks

⇒ 12.5 BTC every 3972 days

⇒ $50000 once every 10.9 years

Mining with mining pool

- Assume pool has ⅙ network hashrate
  - Pool finds every 6th block ≈ 1 per hour
- $5080 / 8760 hrs/yr

≈ $0.58 every hour

Paradox:

- The more secure Bitcoin gets, the greater the appeal for mining pools

IDENTITY IN BITCOIN

- **Identity pairs**
  - Email: address and passwords
  - House: address and keys
  - Bitcoins: public key and private key

- **Identity pairs in the context of currencies**
  - public key for receiving money
  - private key for redeeming money

IDENTITY: PUBLIC AND PRIVATE KEY GENERATION

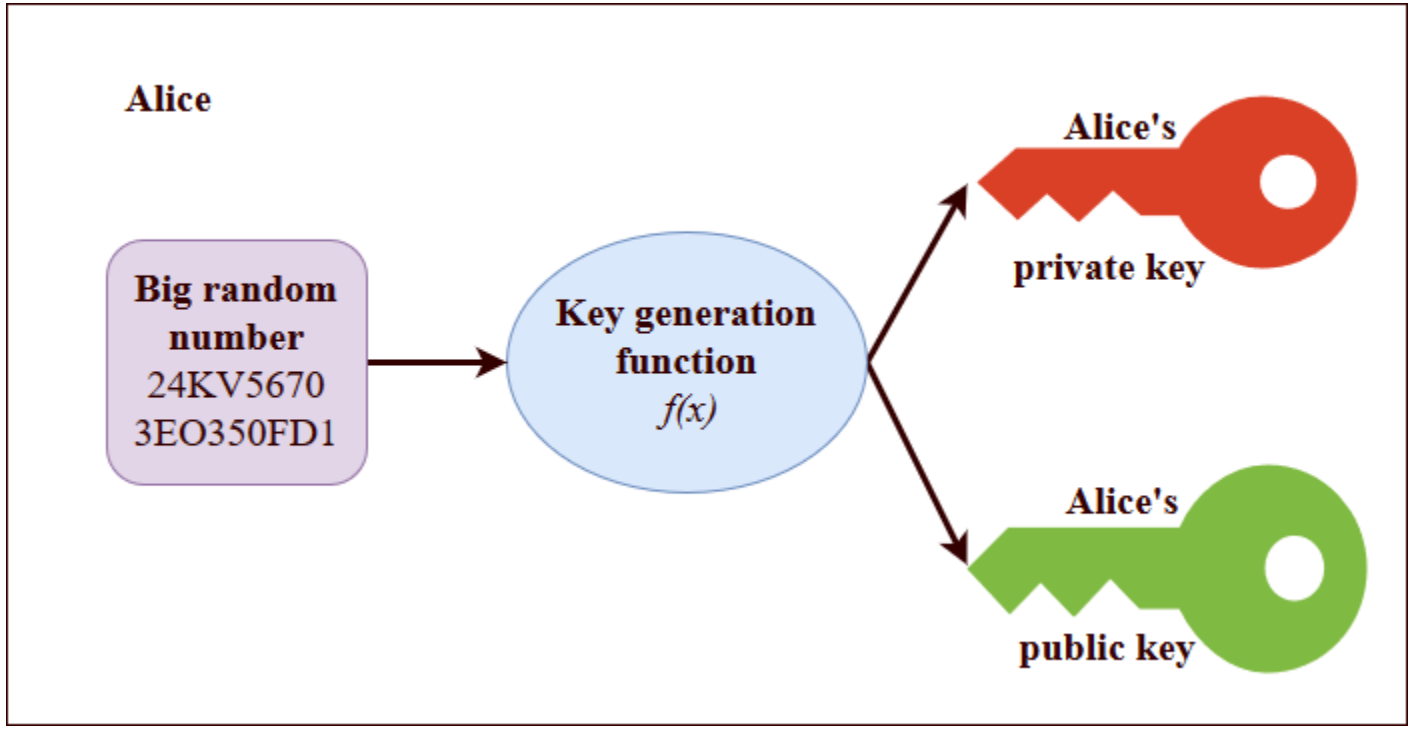
Key Management:

Normal practice is to generate new keys for every transaction you make

Why?

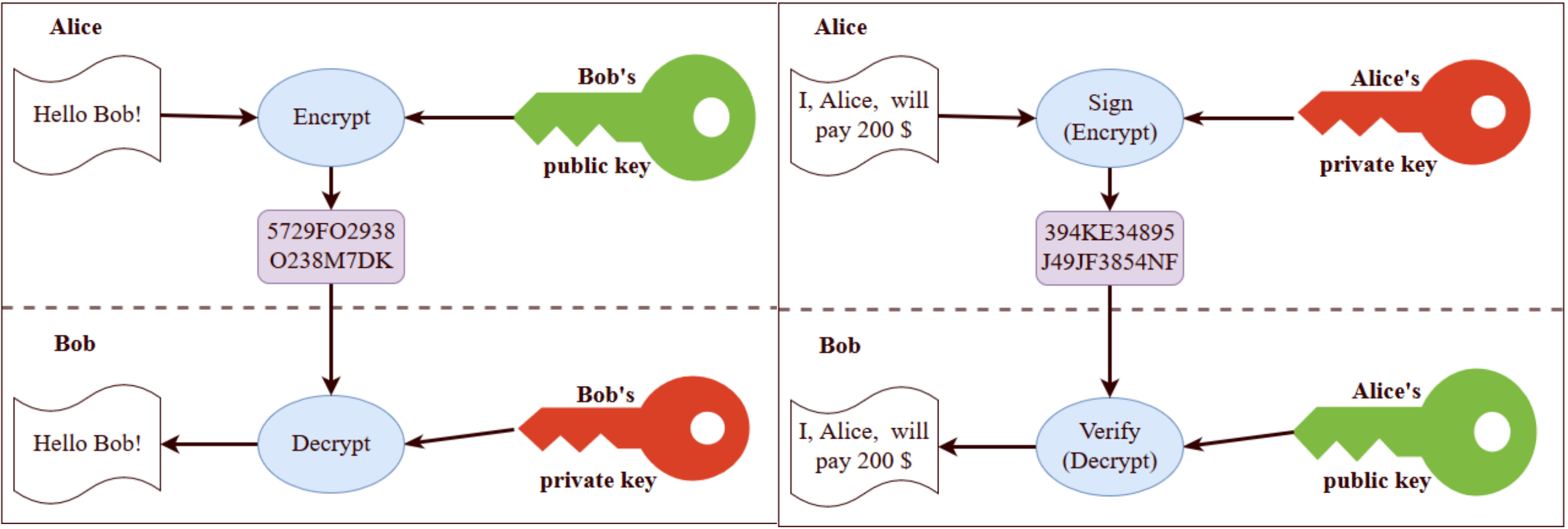
- Privacy
  - Someone shouldn't be able to determine how much bitcoin you own
- Easier to tell who sent you a transaction

Wallet software will do legwork of combining funds from different keys



Public key for receiving

Private key for redeeming



"What if someone guesses my private key?!"

- Bitcoin is hidden in the large amount of public keys
- 2^{160}

  (1,461,501,637,330,902,918,203,684,832,716,283,019,655,932,542,97)

  possible addresses
- Practically impossible for anyone to overlap
  - For reference:
    - Grains of sand on earth: 2^{63}
    - With 2^{63} earths, each with 2^{63} grains of sand: 2^{126} total grains of sand
    - 2^{126} is only 0.0000000058% of 2^{160}
  - Population of world: 7.5 billion in April 2017
    - Every person could have about 2^{127} addresses all to themselves

TRANSACTIONS: VALIDITY

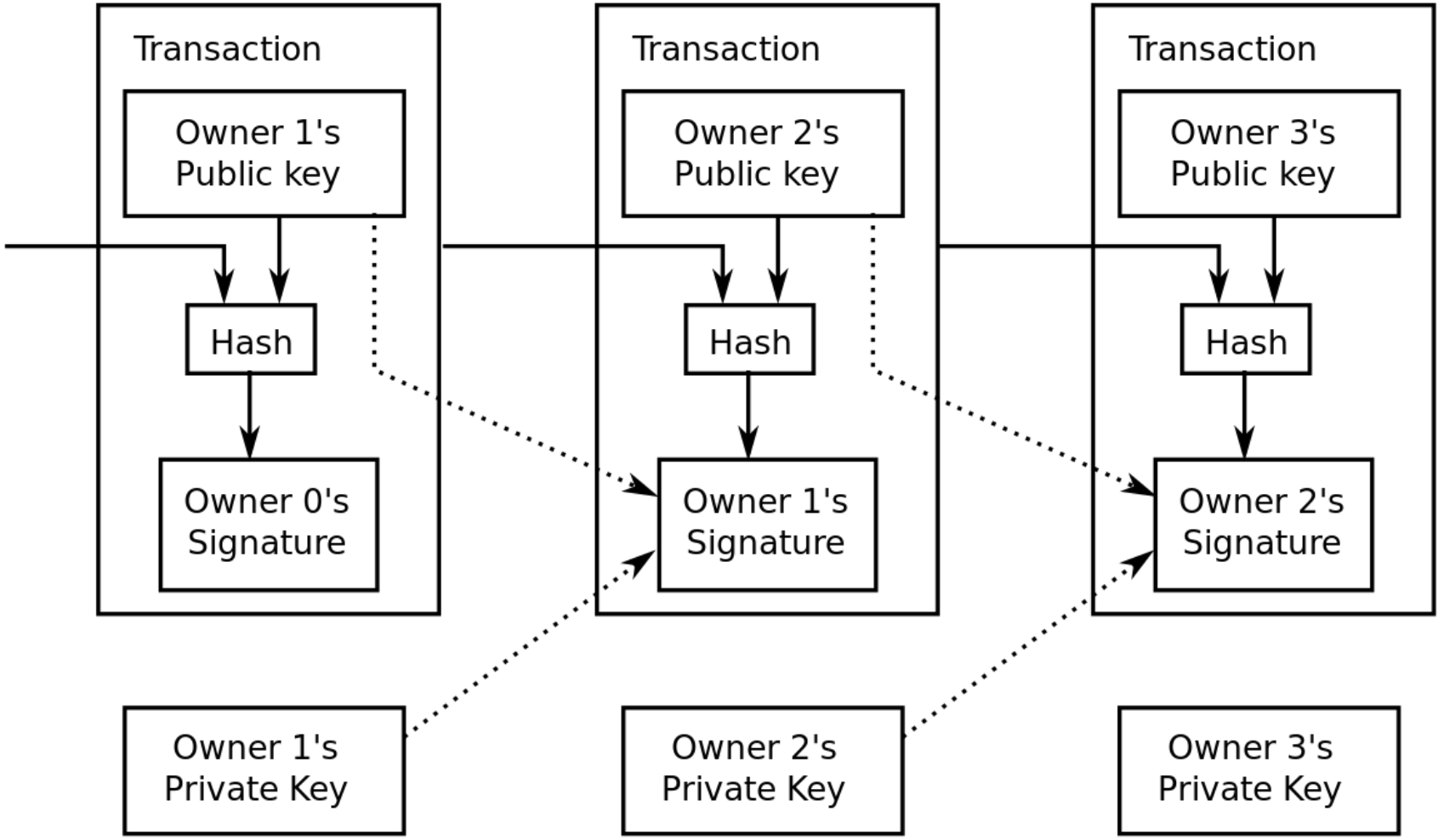
A **transaction** must be **valid**, if all the conditions below are satisfied:

- signature
- funds available
- no other transactions using the same fund

Unspent transaction output (UTXO) model

- to ensure the funds are used only once

TRANSACTIONS (1/2)

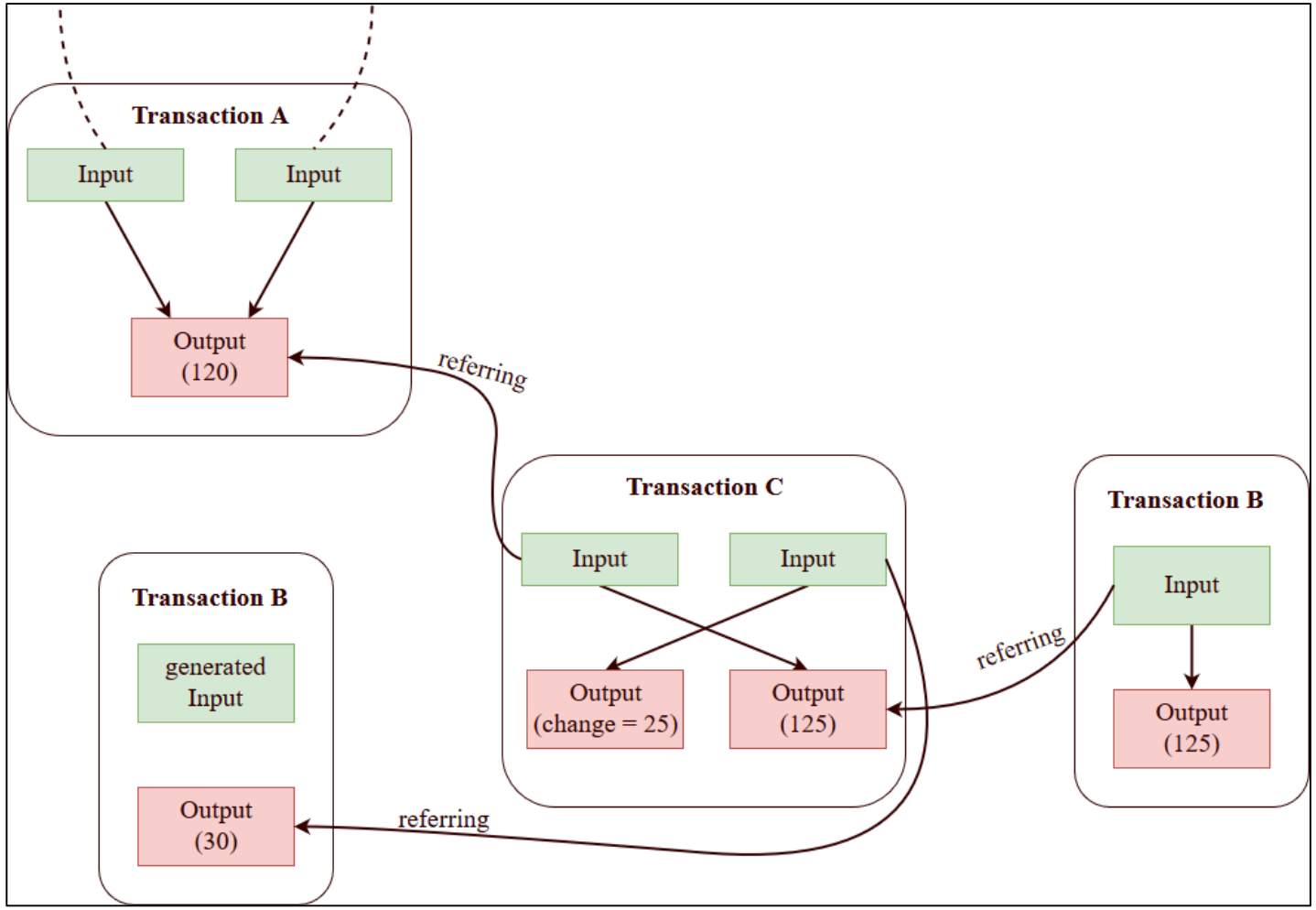


TRANSACTIONS (2/2)

	version	01 00 00 00
	input count	01
Input	**previous output hash (reversed)**	jd 23 i7 jr 85 e9 2m 19
Input	**previous output index**	00 00 00 00
Input	**script length**	
Input	**scriptSig**	script containing signature
Input	**sequence**	ff ff ff ff
	output count	01
Output	**value**	A5 00 00 00 00 00
Output	**script length**	
Output	**scriptPubKey**	script containing destination address
	block lock time	00 00 00 00

TRANSACTION: UTXO

- Instead of keeping all your cash in one chest, each received payment goes into a new piggy bank
- Every time you need to make a transaction, you break one or more piggy banks
- All bitcoins have a „serial number“, the reference number when using UTXOs as inputs for other transactions.



POTENTIAL ATTACKS (1/2)

Sybil attack

- A malicious adversary creates just copies of nodes, all controlled by the same party
  → Bitcoin‘s Solution: Implicit consensus (by randomness) for a block creation

Denial of service attack

- A denies transactions from B‘s address

51% attack

Bitcoin‘s solution → PoW (Bitcoin assumes an honest majority)

POTENTIAL ATTACKS (2/2)

Double-spending attack

- Alice creates two transactions: one to Bob, the other to herself
- Bitcoin‘s solution → make sure to wait for ***n*** confirmation.
    - **Confirmation:** The number of blocks created on top of the block a txn is in.

$$confirmations = block_{depth} - 1$$