CSE446: Blockchain & Cryptocurrencies

Lecture – 18: Blockchain Properties, Misconceptions & Limitations

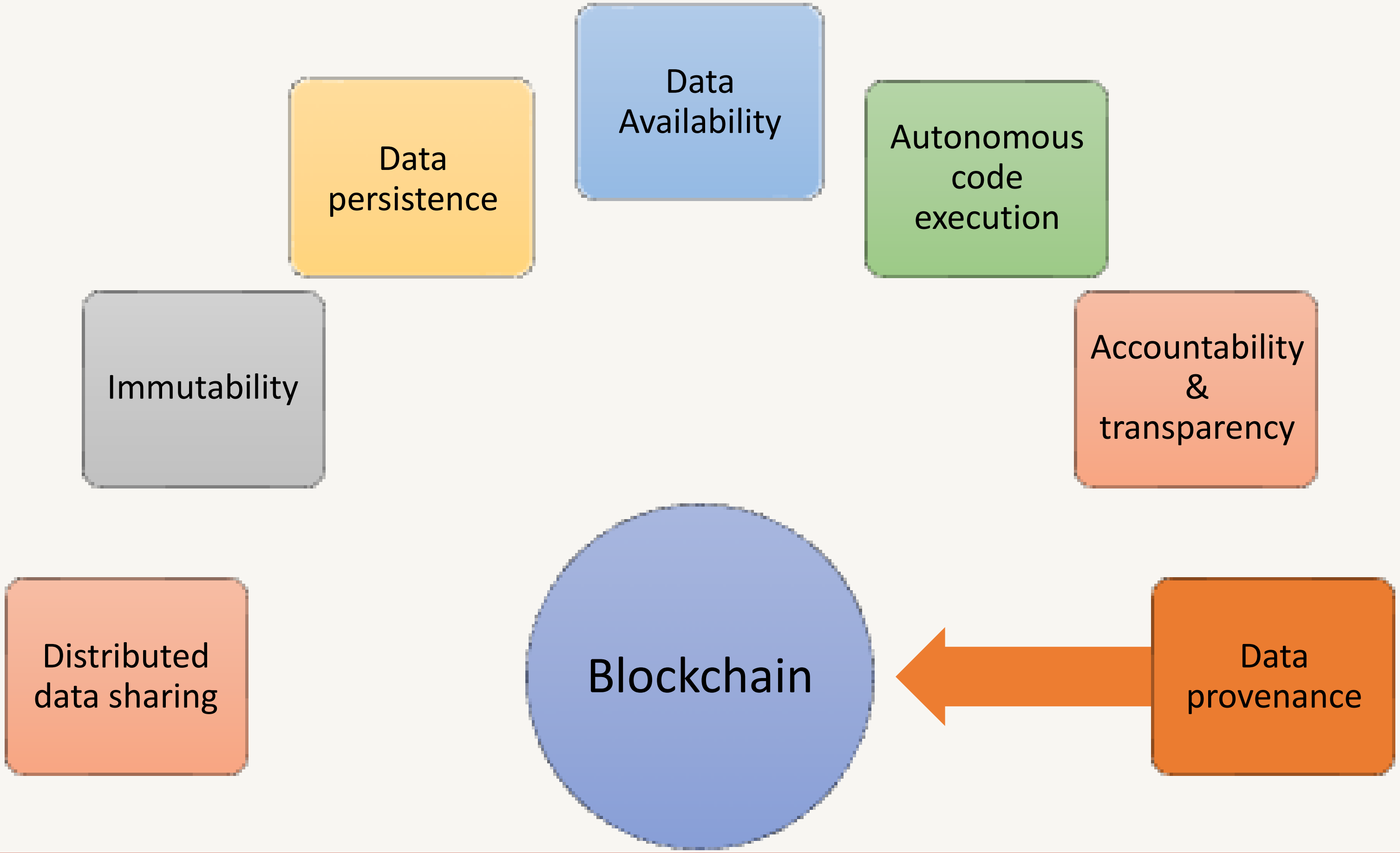
BRAC UNIVERSITY

Inspiring Excellence

Agenda

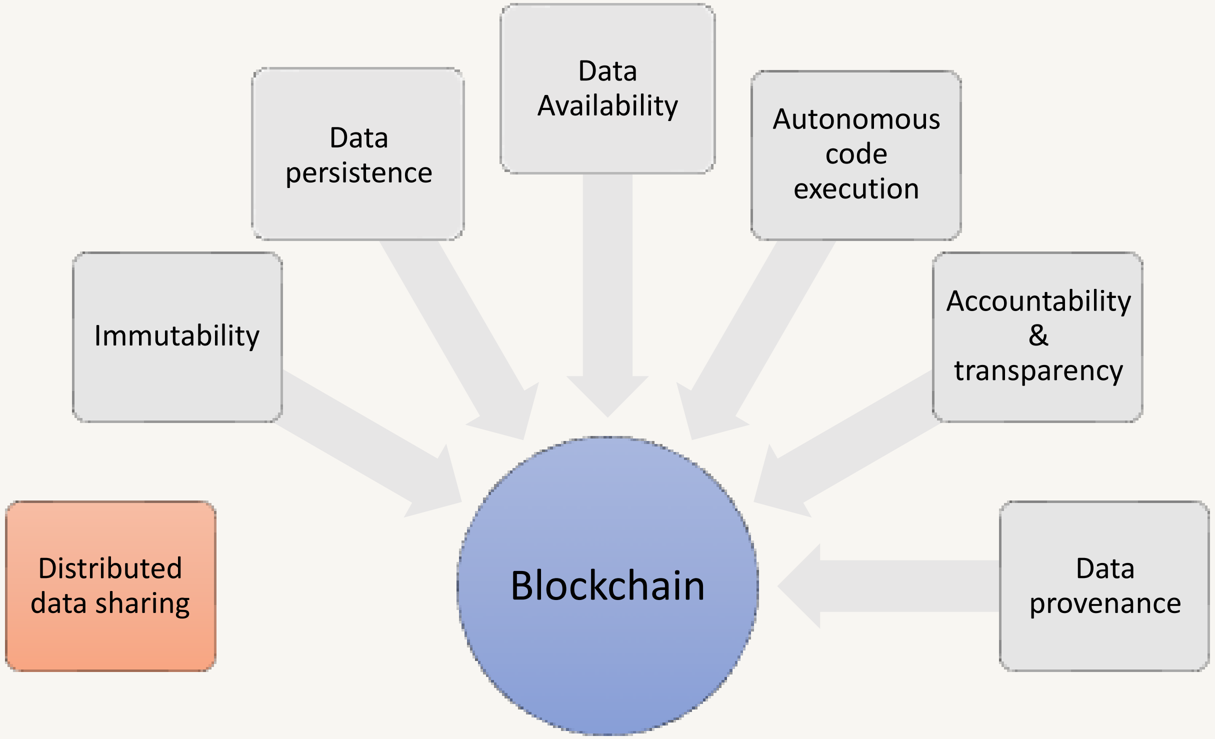
- Blockchain Properties
- Blockchain Misconceptions
- Blockchain Limitations
- Blockchain feasibility
- Attacks on blockchain

Blockchain properties



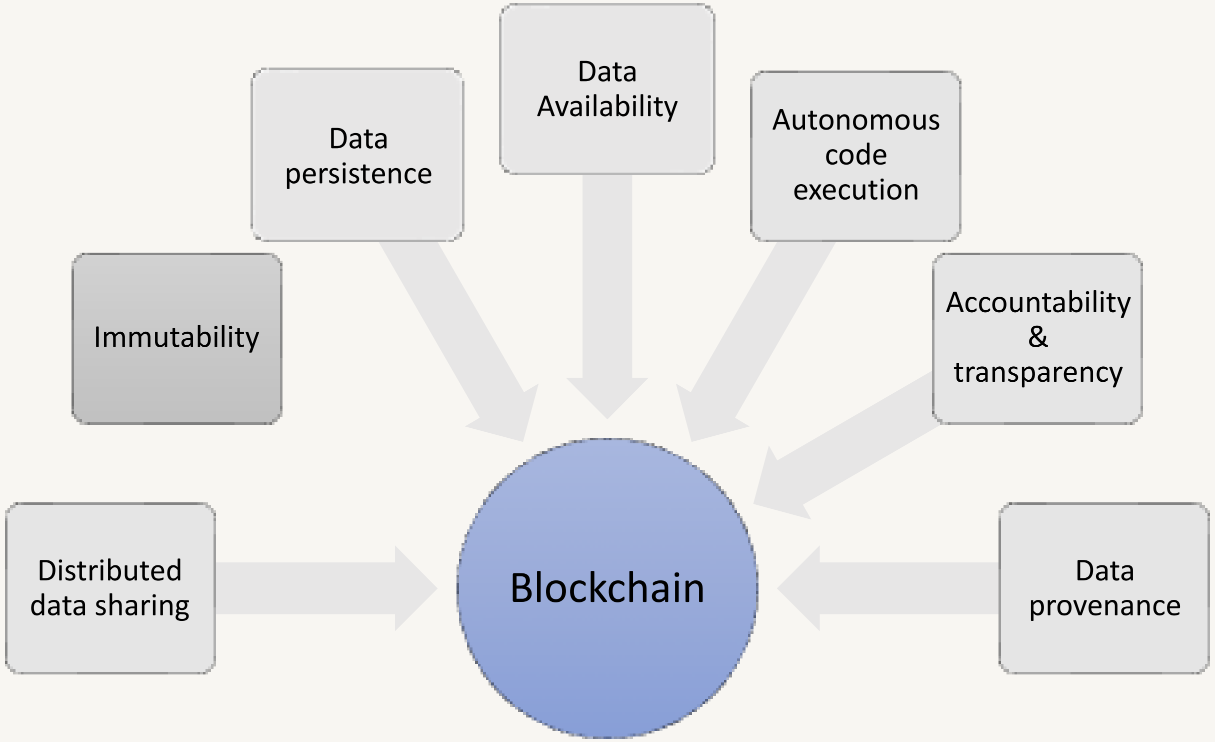
Distributed Data Sharing

- Blockchain data is distributed across multiple nodes
- The protocol ensures that data inserted in a particular node gets synced across all nodes in a timely fashion



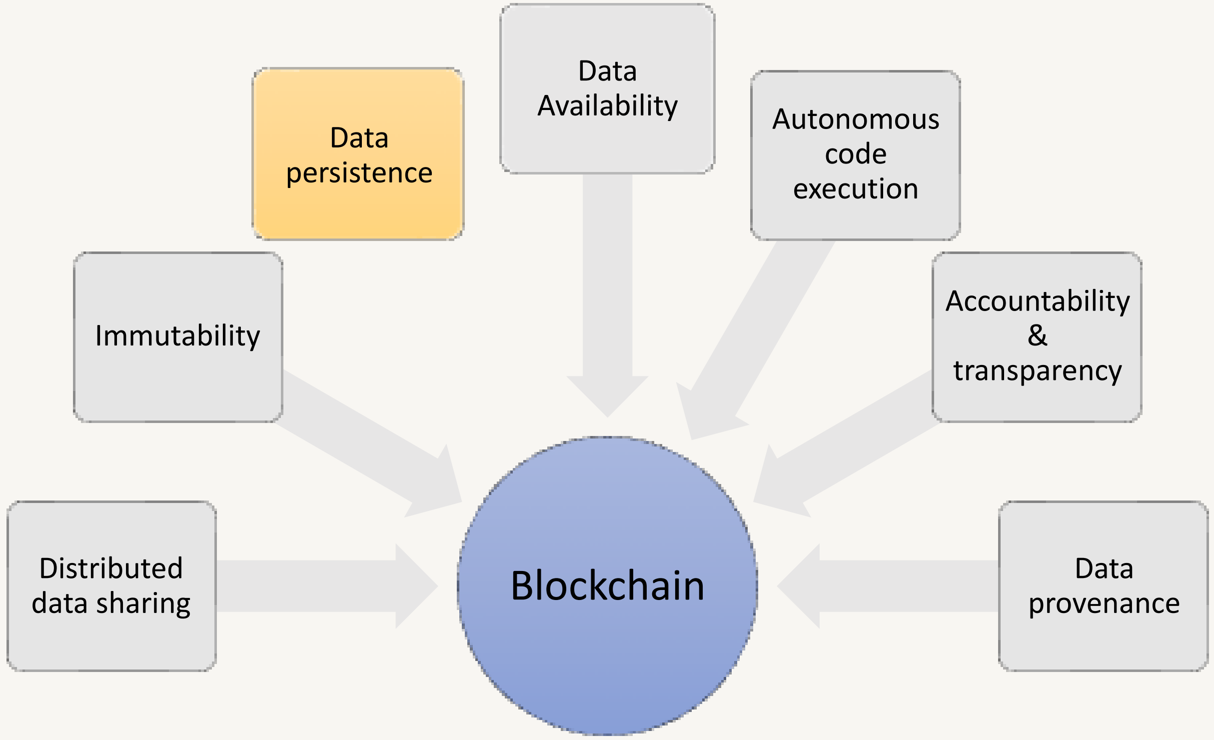
Immutability

- Data and code immutability in blockchain emerges from the fact that to change data/code inserted in a previous block, an attacker must posses either
  - significant computational power, in case of a public blockhain
  - or compromise the majority of nodes in any type of blockchain



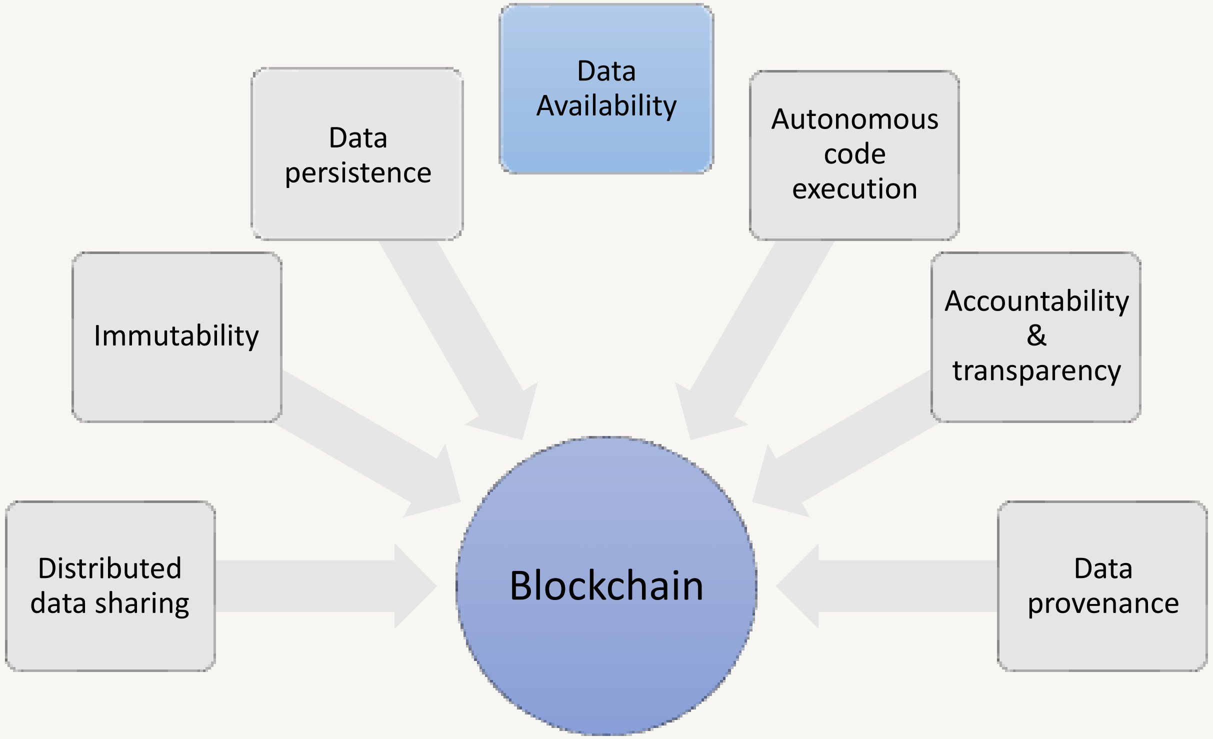
Data persistence

- Being a distributed system implies that data in a blockchain will persist as long as there are enough nodes to execute the protocol in a secure way



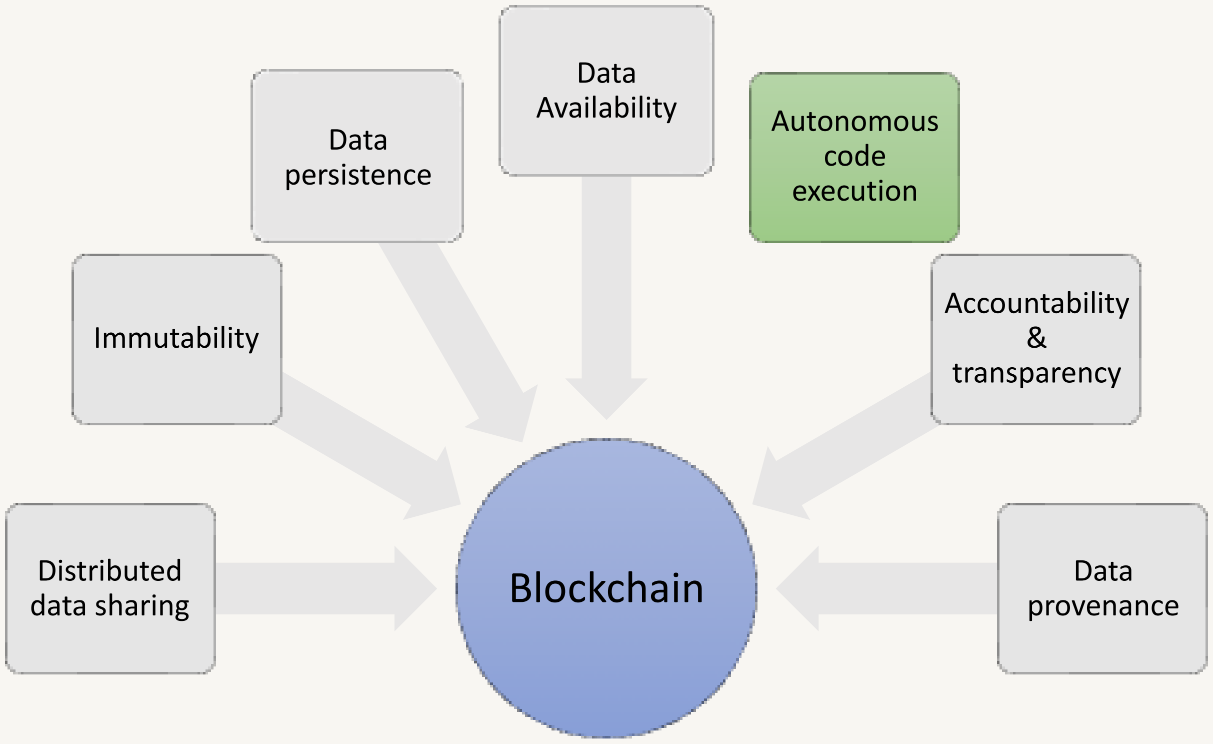
Data availability

- Data in blockchain are always available
- Even when a particular node is offline, data can be retrieved from another node in the blockchain



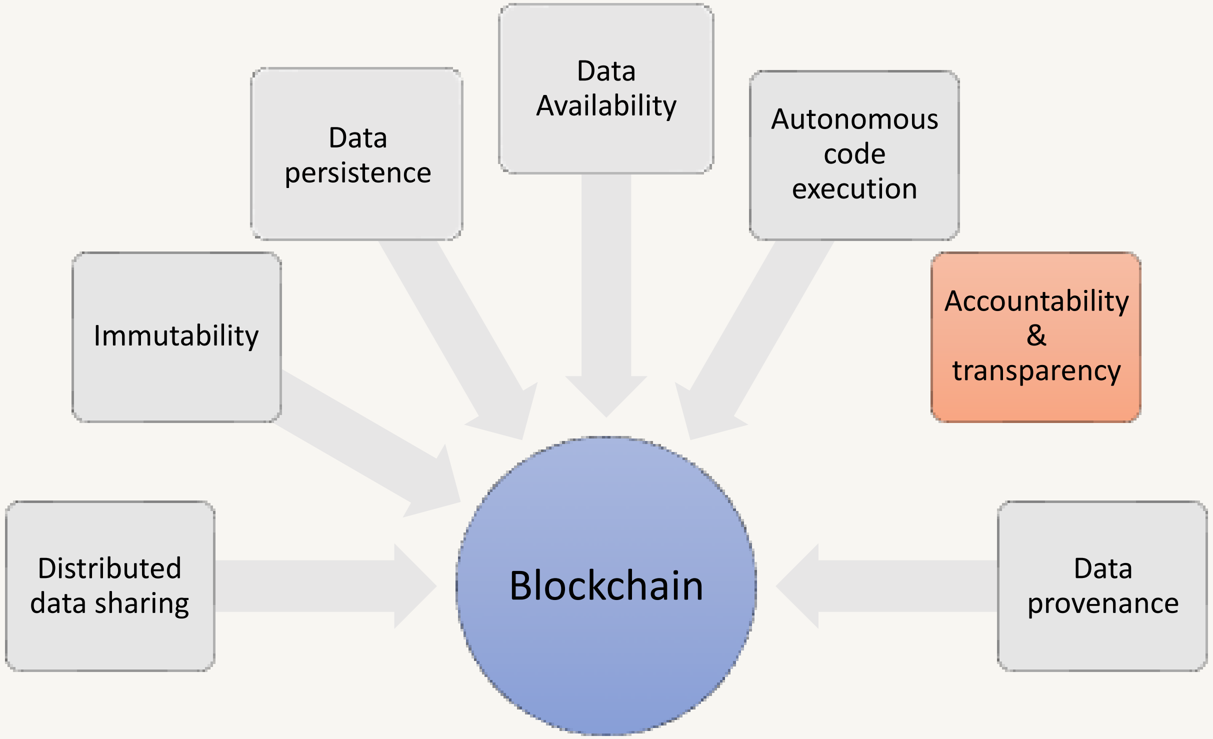
Autonomous code execution

- A smart-contract will facilitate autonomous code execution without a single point of failure
- It does not require any human intervention
  - anyone can submit a transaction to execute a code
- For any private blockchain system, anyone authorised can execute a code



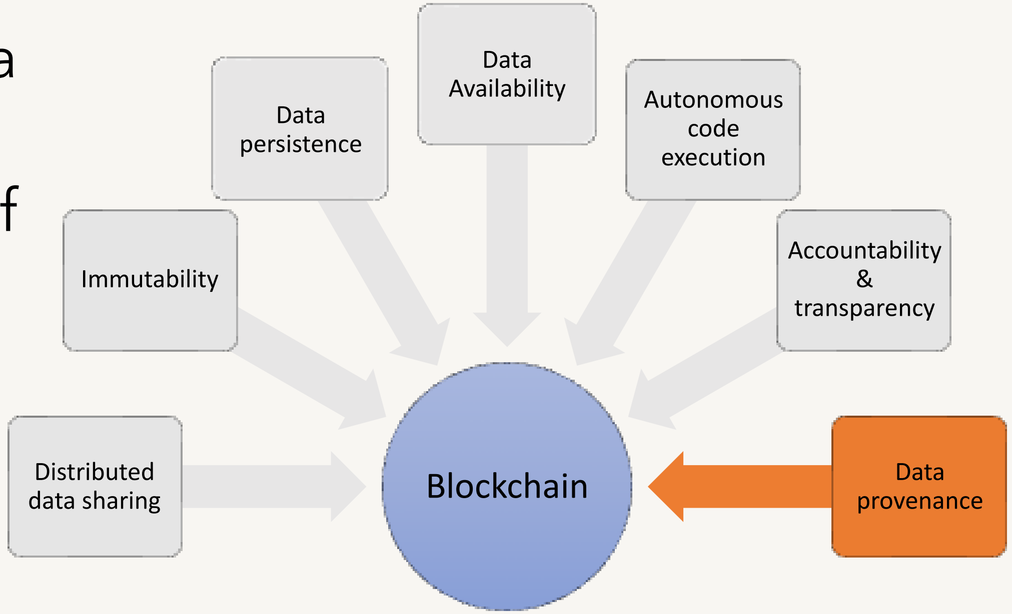
Accountability and transparency

- All authorised entities can verify each single transaction which can ensure accountability and transparency



Data provenance

- The term "data provenance" refers to a record trail that accounts for the origin of a piece of data (in a database, document or repository) together with an explanation of how and why it got to the present place
- Data in a blockchain can only be stored with a signed transaction
- Blockchain also stores the transactions which might have changed the data
- Both of these ensure data provenance



Blockchain misconception

DATA IMMUTABILITY

LARGE-SCALE DATA STORAGE

DATA INTEGRITY

DATA ENCRYPTION

POWER CONSUMPTION

Data immutability misconception

- Blockchain data can be never be changed
- This is true for transaction/blockchain data which are immutable
- However, smart-contract data can be changed as required
  - Remember we could change different variable values in a smart-contract
  - However, how such data is changed is recorded in the blockchain and hence, is immutable

DATA IMMUTABILITY

Large-scale data storage misconception

- Blockchain provides integrity of data and hence, users are tempted to store large amount of data in blockchain to ensure integrity
- Performance of any database in terms of data access rate is much better than that of any blockchain system
- Also storing a large amount of data in a public blockchain is costly
- Thus, it is advisable to store as minimum data as possible in the blockchain

LARGE-SCALE DATA STORAGE

Data integrity misconception

- People think blockchain can support data integrity for any type of data
- However, it must be remembered that a blockchain system is essentially a "*Garbage-in-garbage-out*" system
- A corrupted data will be stored and remain as corrupted
- It can guarantee the integrity of data only after it is stored in the blockchain

DATA INTEGRITY

Data encryption misconception

- Many believe that a blockchain provides data encryption by default
- A blockchain system strongly depends on cryptographic mechanisms, such as digital signature and cryptographic hash, to function
- Digital signature is used for data provenance while a cryptographic hash is used to ensure data integrity
- In a blockchain system, data encryption is not provided

DATA ENCRYPTION

Power consumption misconception

- Many believe that every blockchain system consumes a huge amount of power
- However, the reality is that only public blockchain systems which utilise PoW or similar consensus algorithms consume huge electricity
- Public blockchain systems with PoS or DPoS consume significantly less electricity
- The power consumption of any private blockchain system will be comparable to any existing system

POWER CONSUMPTION

Blockchain limitations

- Blockchain bloating
- Blockchain scaling
- Security vs Decentralisation vs Scaling
- Power consumption (already covered)
- Code immutability
- Usability
- Associated expense

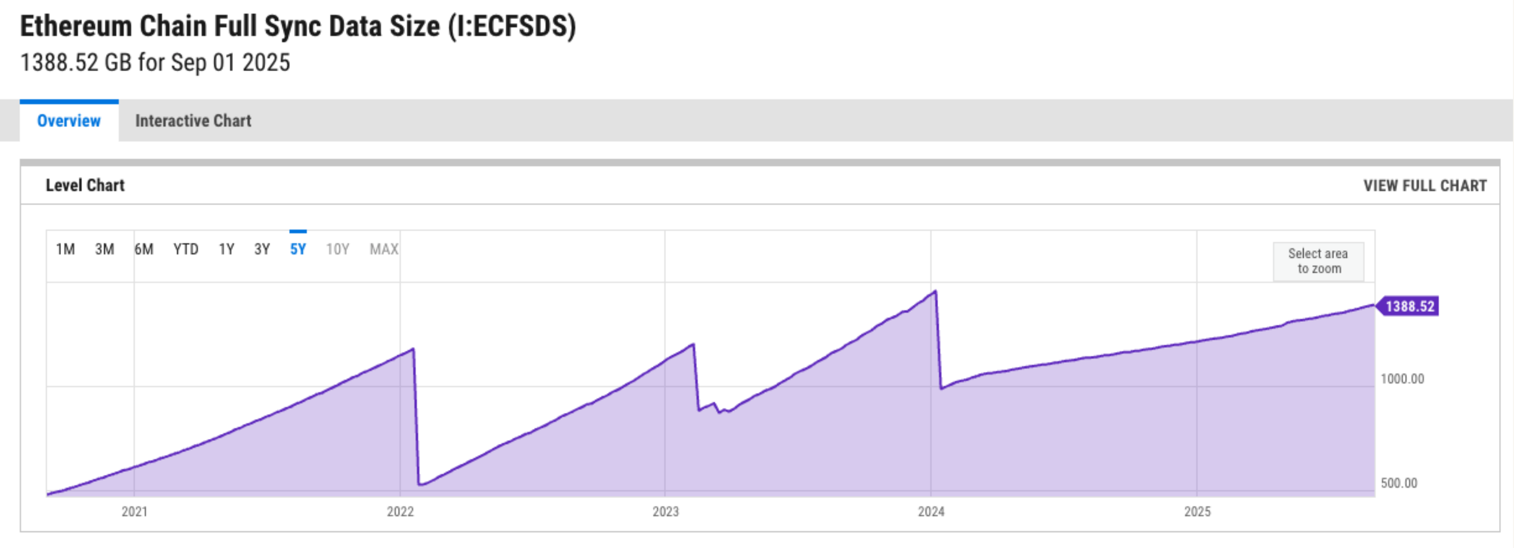
Blockchain bloating

- Blockchain being an add-only distributed database, its size keeps increasing
- Bitcoin size is currently nearly 700 GB and increasing

https://www.blockchain.com/explorer/charts/blocks-size

Blockchain bloating

- Ethereum size is currently >1TB and increasing
- What will happen in 20/30/50 years time?
- Blockchain bloating would also increase data processing time
    - Finding a particular UTXO and so on



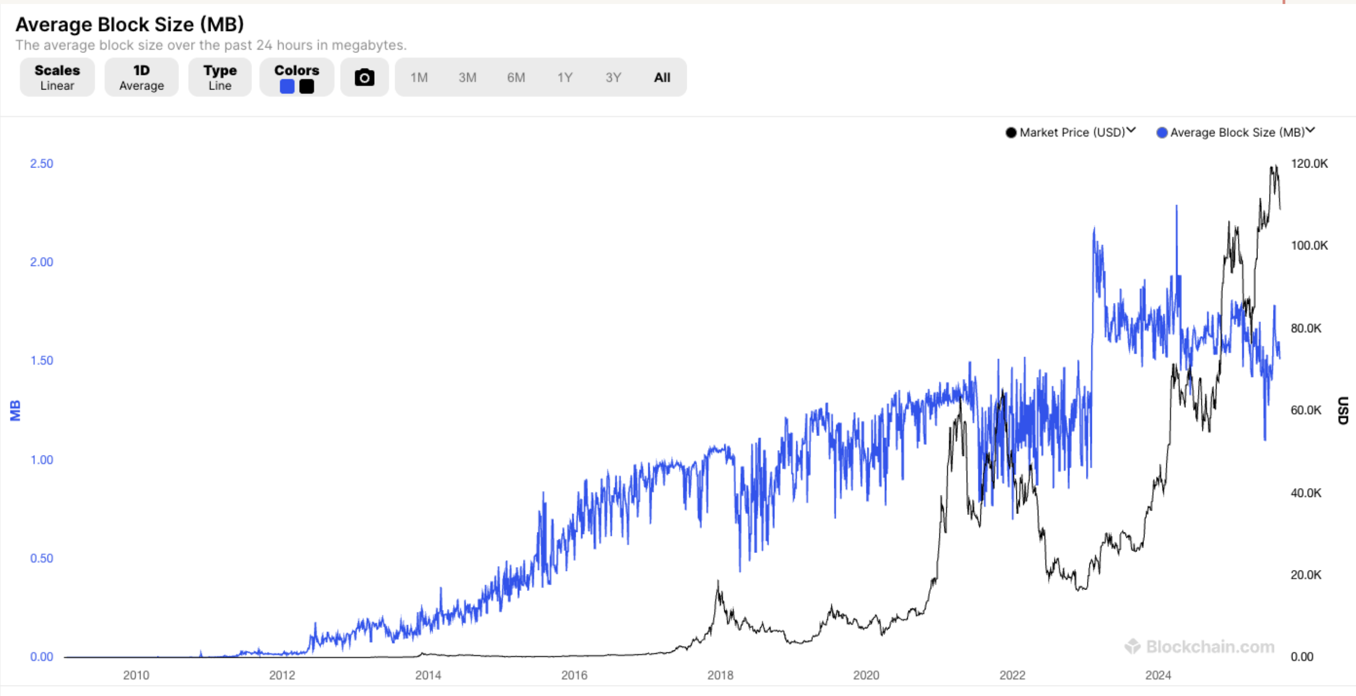
https://ycharts.com/indicators/ethereum_chain_full_sync_data_size

Blockchain scaling

- The blockchain scalability problem refers to the limited capability of the blockchain network to handle large amounts of transactions on its platform in a short span of time
- This limited capability is due to two reasons:
  - Limitations in block size
  - Limited TPS (transaction per second)

Blockchain scaling

- Bitcoin had a limitation of 1MB block size, currently around 2MB
- That means the transactions need to be selected in such a way so that total block size including header is around 2 MB
- Also, the TPS in Bitcoin is around 3-7
- Both of these imply that Bitcoin finds it difficult to handle a large number of transactions in a short period of time

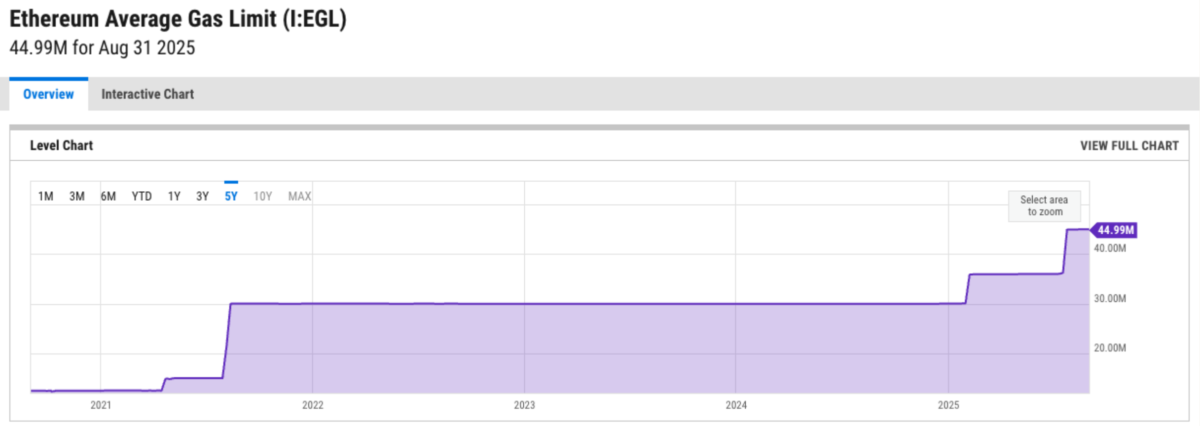


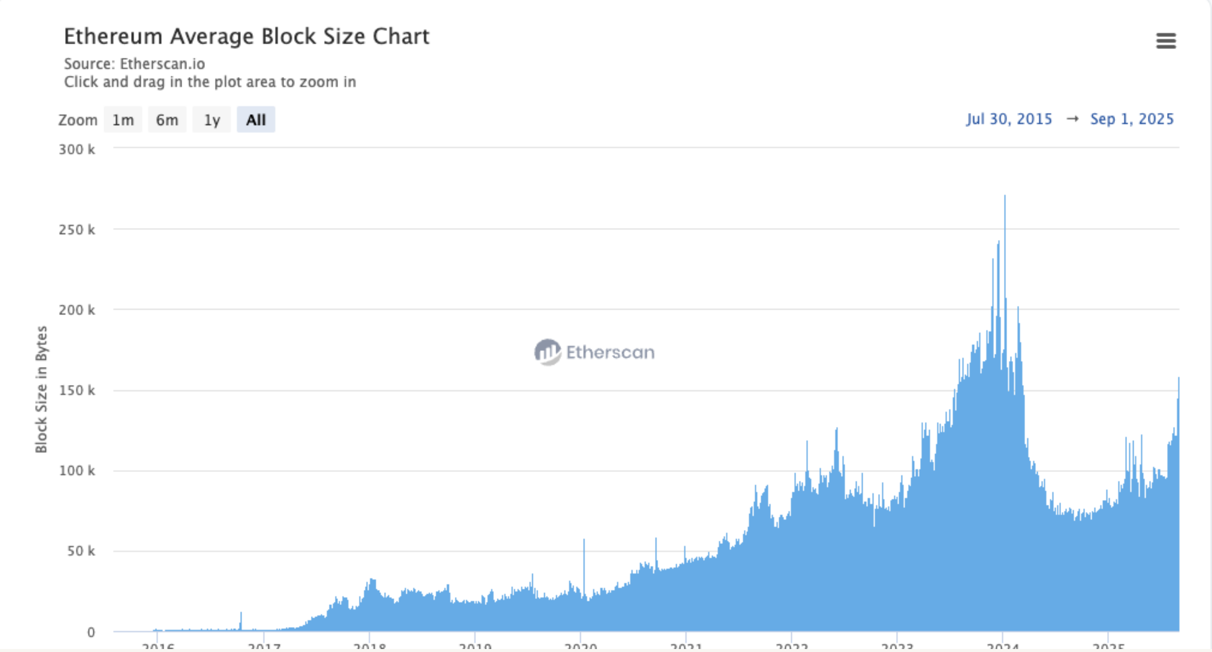
https://www.blockchain.com/explorer/charts/avg-block-size

Blockchain scaling

- In Ethereum, each block has a target size of around 45 million gas
- The size of blocks will increase or decrease in accordance with network demands
- Currently, it results in 1.5MB block size in average
- TPS in Ethereum is around 18

https://ycharts.com/indicators/ethereum_average_gas_limit

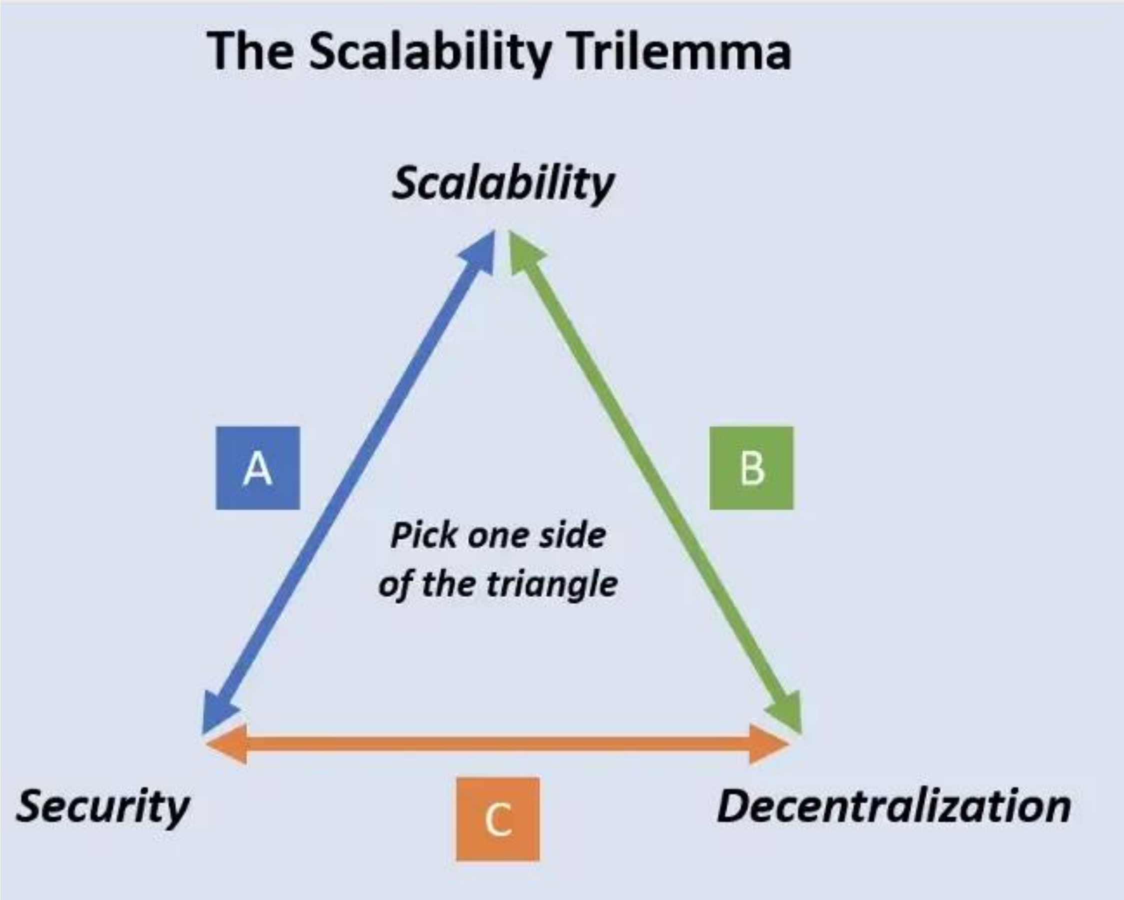




https://etherscan.io/chart/blocksize

Security vs Decentralisation vs Scaling

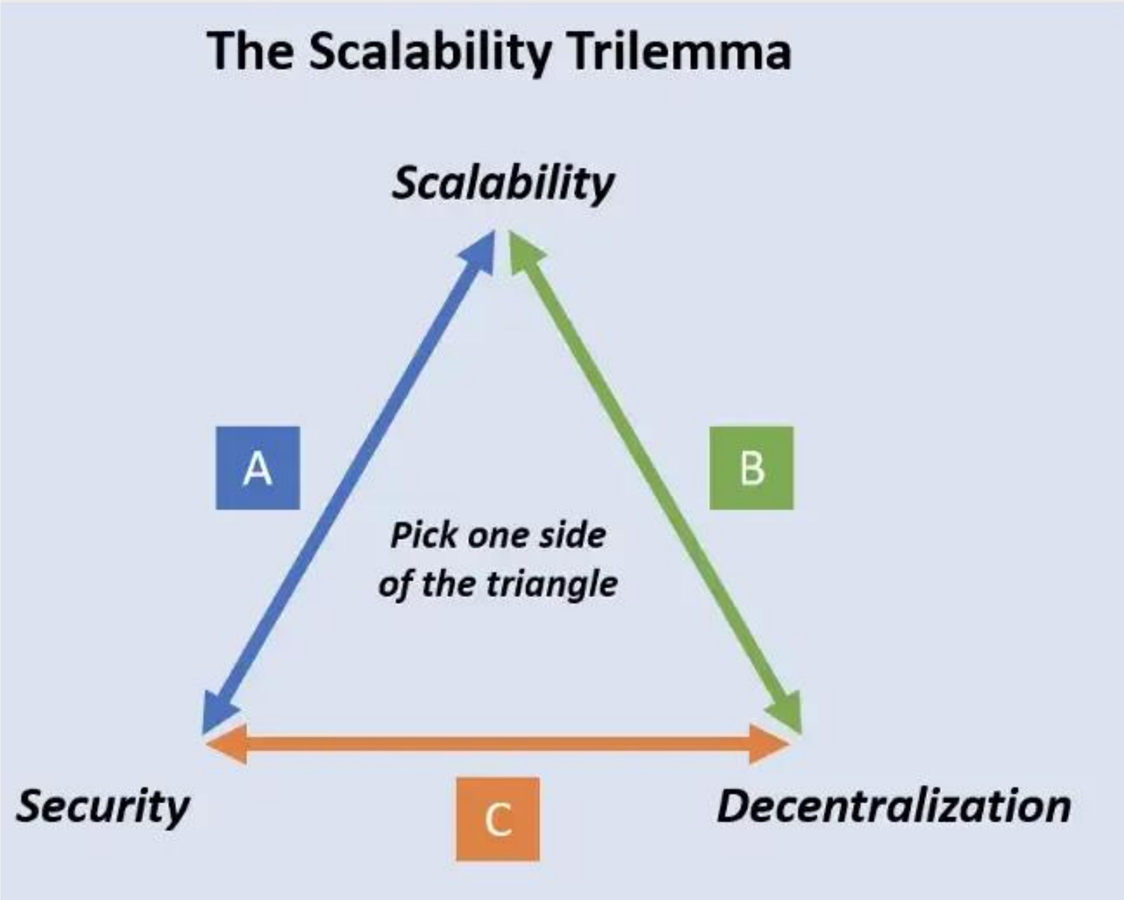
- Blockchain scalability trilemma was introduced by Vitalik
  - "trilemma" is a situation where you can only get two out of three desirable outcomes
- The trilemma is we can only choose any of the two properties from these three
  - Security and Scalability, **Security and Decentralisation**, Decentralisation and Scalability



https://miro.medium.com/max/720/1*oYKf61tE6dQPPYft2tqWow.webp

Security vs Decentralisation vs Scaling

- Decentralisation implies not utilising a single central system
- This will also increase the security
  - It is not a single point of failure anymore
- However, this might reduce the TPS
  - As it would require slightly more time to reach a consensus
  - Remember, more people mean more arguments and difficult to reach a consensus



https://miro.medium.com/max/720/1*oYKf61tE6dQPPYft2tqWow.webp

Code immutability

- Once a smart-contract is deployed it becomes immutable
- This has huge advantage; however, it also introduces limitations
- If there is a bug in a smart-contract it cannot be rectified
- The error needs to be fixed off-chain and then re-deployed in the network
  - This will result in a new contract address
- The Dapp then needs to be updated with the new contract address

Usability

- Remember using Metamask setup?
- Need to remember the password or safely store the mnemonic words (passphrase)
  - If you forget your password and don't safely backup your passphrase then you cannot recover your wallet or funds in it
- Think about, how did you find it: easy or difficult?
- Then think about the general people and how they would find it to use such a complex process?
- Researchers have found strong usability issues with Blockchain wallets

Associated expense

- It takes considerable investment to join the mining and staking process
- Storage and computation costs crypto-currency (eth)
- In a 2018 estimation:
  - When 1 ETH = 200 USD
  - 1 KB required 2 USD
  - 1 MB would cost around 2000 USD
- But now, 1 ETH = 3000 USD
  - It is very difficult to predict, how much it might cost
  - One option is to try via Ganache or in the test network

Blockchain feasibility

DECENTRALI
SATION

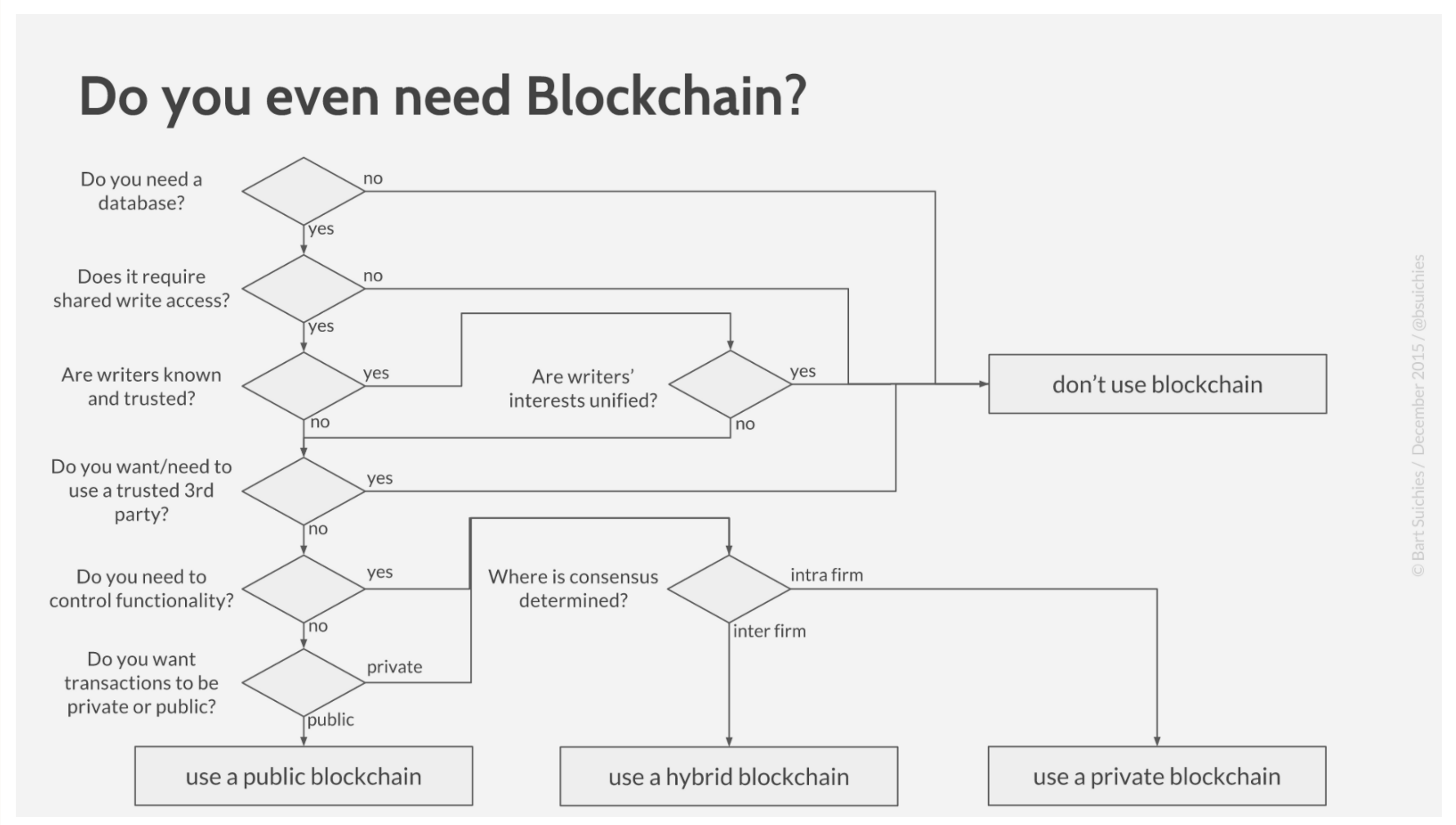
DISINTERME
DIATION

P2P VALUE
TRANSFER

DATA/CODE
IMMUTABILITY

DISTRIBUTED
DATA SHARING

Blockchain feasibility



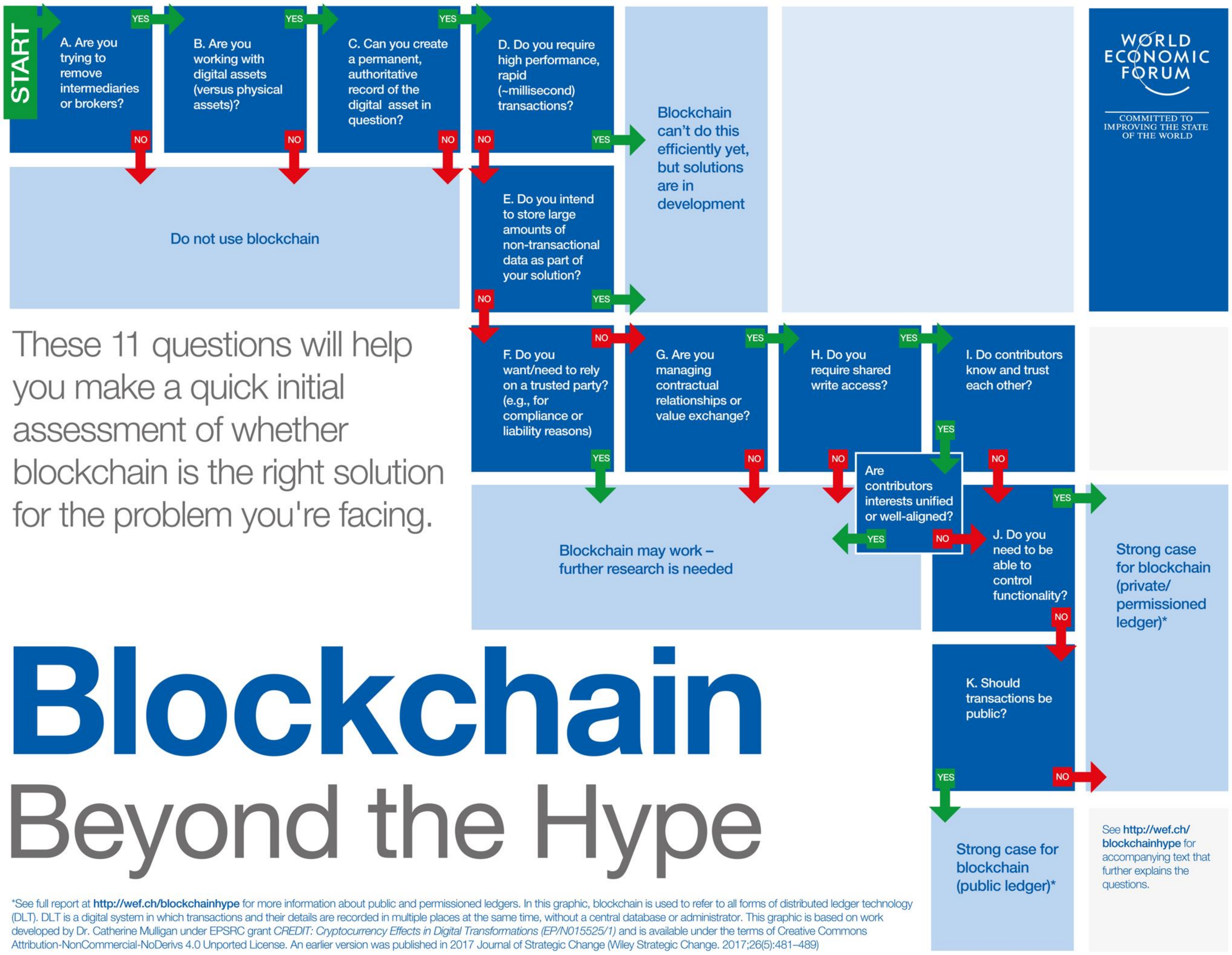
Bart Suichies model

https://medium.com/@bsuichies/why-blockchain-must-die-in-2016-e992774c03b4

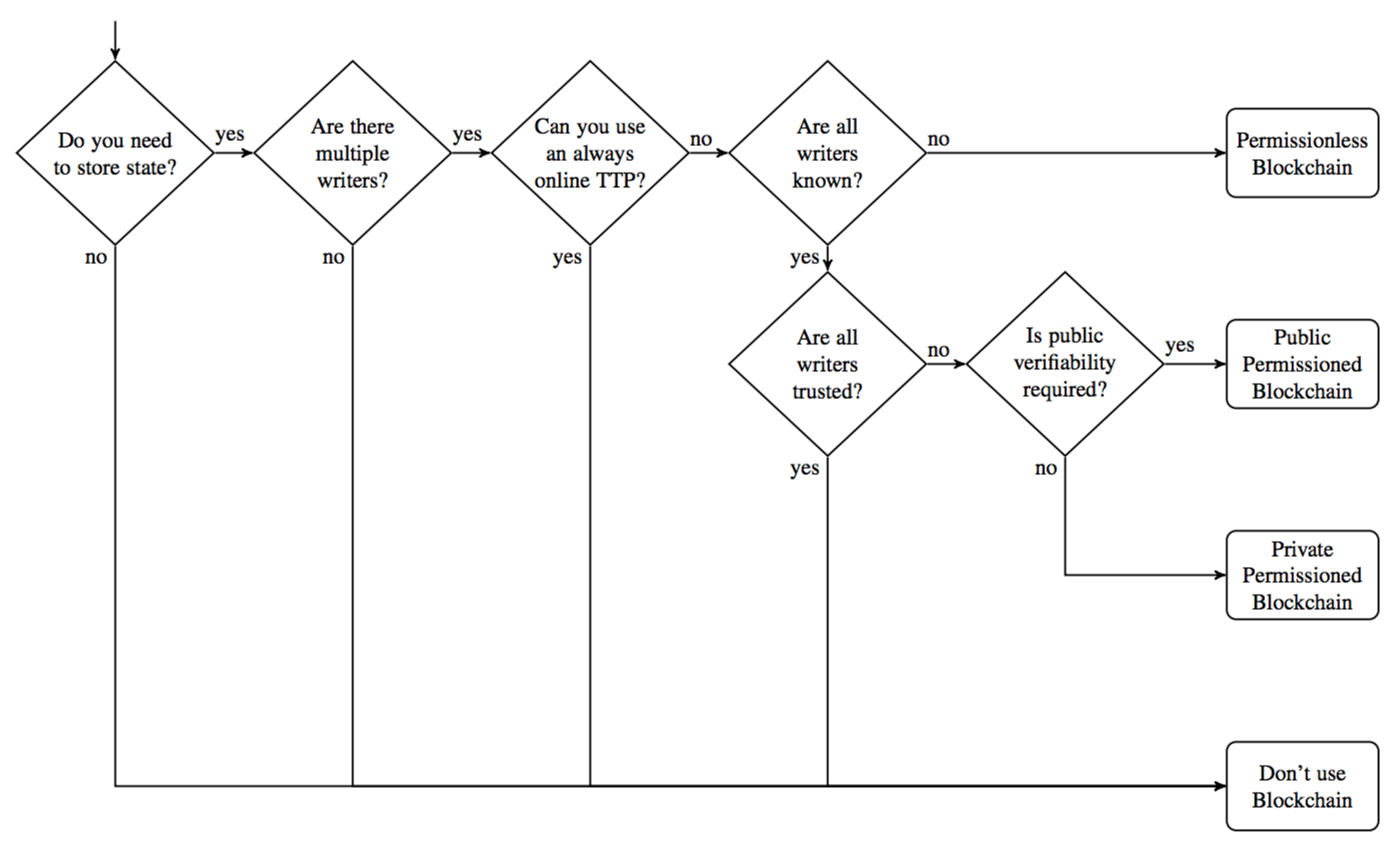
Blockchain

WEF Model

https://www.weforum.org/agenda/2018/04/questions-blockchain-toolkit-right-for-business



Blockchain feasibility



Model by Karl Wüst & Arthur Gervais

Wüst, Karl, and Arthur Gervais. "Do you need a Blockchain?." IACR Cryptology ePrint Archive 2017 (2017): 375.

Attacks on Blockchain

- Transaction censoring attack
- 51% attack
- Double-spending attack
- Selfish mining attack (Block withholding attack)
- Sybil attack
- DDoS attack

Transaction censoring attack

- Blocking (censoring) a transaction from a certain address (people)
- Malicious validating (full) nodes:
  - As long as there are majority of honest nodes (>50%), the transactions will be propagated
  - Mind you, a malicious node cannot affect the blockchain in this way
- Malicious mining nodes:
  - If malicious mining nodes censor transactions, they would still be included in a block mined by an honest node

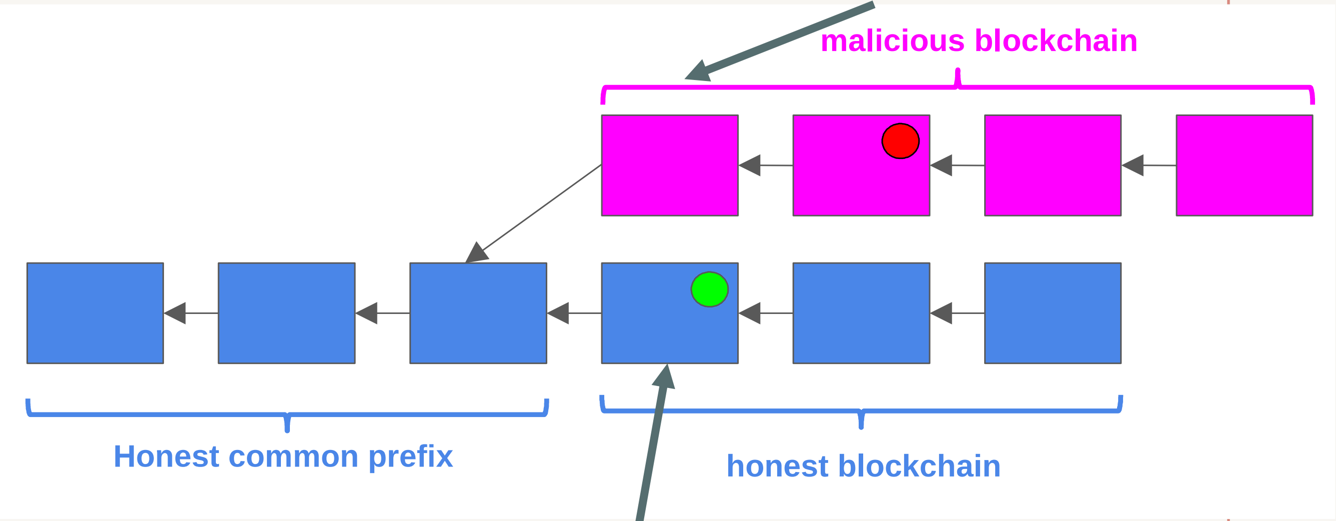
51% attack

- A group of malicious nodes can collude to launch the infamous 51% attack
- It happens in PoW-based blockchains, if a single miner’s hashing power accounts for more than 50% of the total hashing power of the entire blockchain
- In PoS blockchain, 51% attack may also occur if the number of coins owned by a single miner is more than 50% of the total blockchain
- Controlling a majority (51%) can cause a deliberate "fork" in the blockchain
  - A fork is where the attacker causes previously confirmed blocks to be invalidated by forking below them and re-converging on an alternate chain
  - With sufficient power or tokens, attackers can generate blocks at a faster rate than honest miners
  - This will result in the invalid chain to be longer than the shorter chain, ultimately becoming the main chain

Double-spending attack

- Double spending means the act of using the same currency more than once
- Double-spending one's own transactions is profitable
  - if by invalidating a transaction the attacker can get an irreversible exchange payment or product without paying for it
- To deter this attack, wait for the finality
  - For bitcoin it is 6 confirmations ~ 1 hour

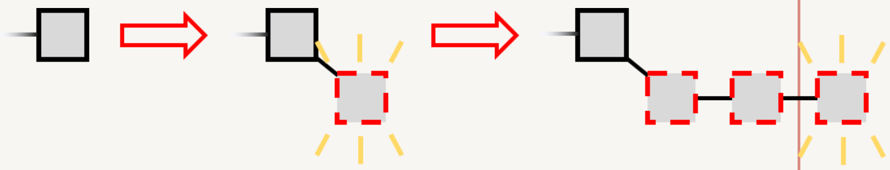
After receiving the good, launches the double spending attack here

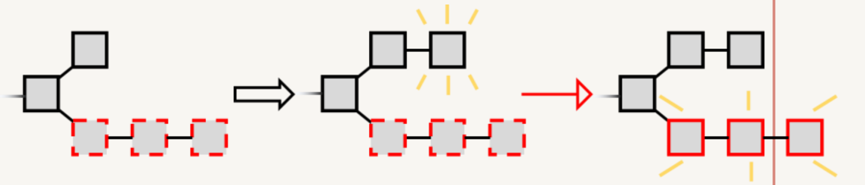


Attacker pays here to buy something

Selfish mining (block withholding) attack

- An attack to PoW blockchain
- The attacker (selfish miner) privately generates valid blocks and extends their own chains secretly, forming a secret branch
- The selfish miner continues to extend her secret branch until the public chain is one step behind
- Then she publishes her secret chain
- Since the secret chain is longer, the other parties consider it the main chain, so now everyone is following the selfish miner's blocks
- The blocks generated by the other miners are ignored
- The selfish miner can reap rewards for multiple blocks together

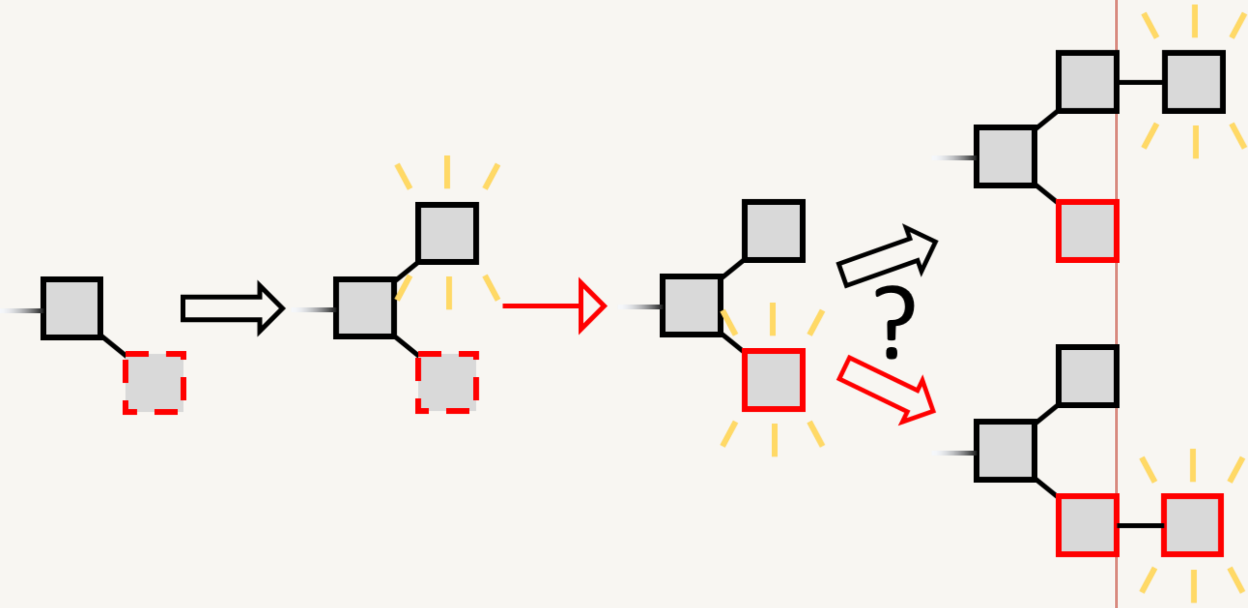




https://decentralizedthoughts.github.io/2020-02-26-selfish-mining/

Selfish mining attack

- But there is a caveat to this strategy - when first forming her secret chain, the selfish miner takes a risk
- If she generated the first secret block and then another miner generated a block at the same time
  - it will be a race between two branches, and it is not guaranteed if the selfish miner would win



https://decentralizedthoughts.github.io/2020-02-26-selfish-mining/

Sybil attack

- In a Sybil attack, an attacker can duplicate her identity as required in order to achieve illicit advantages
- This can be as simple as one person creating multiple social media accounts
- The word "Sybil" in the name comes from a case study about a woman named Sybil Dorsett, who was treated for Dissociative Identity Disorder – also called Multiple Personality Disorder
- Within a blockchain system, a sybil attack implicates the scenario when an adversary can create/control as many nodes as required within the underlying P2P network to exert influence on the distributed consensus algorithm
    - Having more nodes could improve her probability to solve the PoW algorithm or to be selected as a leader to propose a block in a PoS system

Sybil attack

- Attackers may be able to out-vote the honest nodes on the network if they create enough fake identities (or Sybil identities)
- They can then refuse to receive or transmit blocks, effectively blocking other users from a network
- In really large-scale Sybil attacks, where the attackers manage to control the majority of the network computing power or hash rate, they can carry out a 51% attack
- In such cases, they may change the ordering of transactions, and prevent transactions from being confirmed
- They may even reverse transactions that they made while in control, which can lead to double spending

https://academy.binance.com/en/articles/sybil-attacks-explained

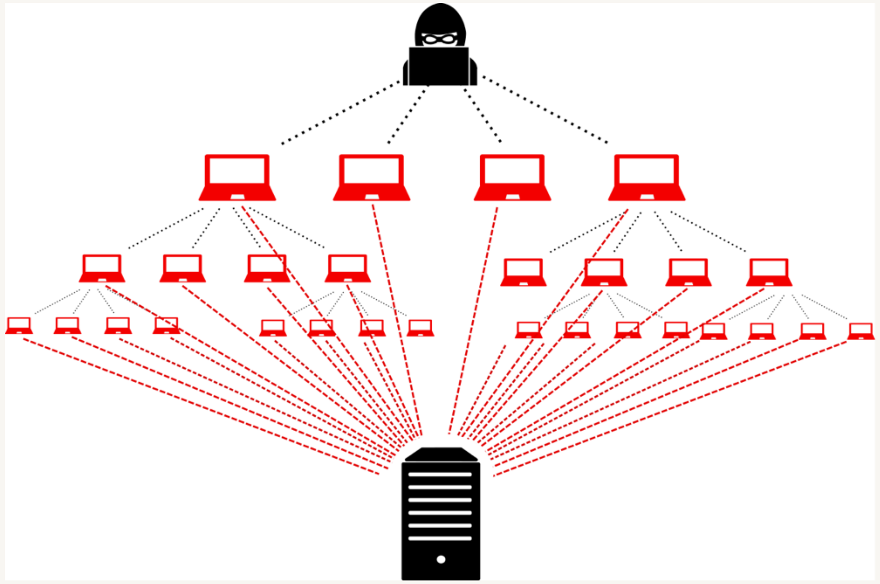
Sybil attack

- Consensus algorithms must be designed in such a way that they can deter any Sybil attack
- These consensus algorithms don't actually prevent Sybil attacks, they just make it very impractical for an attacker to successfully carry out a Sybil attack
- For example, when using PoW algorithms, the probability of creating a valid block is proportional to the total hashing power of all (sybil) nodes
- That means attackers actually need to own the computer power required to create a new block, which makes it very difficult and costly for an attacker to do
- Since mining is so intensive, miners have a very strong incentive to keep mining honestly, instead of attempting a Sybil attack

https://academy.binance.com/en/articles/sybil-attacks-explained

DDoS attack

- DDoS stands for Distributed Denial-of-Service attack
- The attack is distributed over a large network of compromised endpoints
- This large network is called a botnet, which resembles an army under the command of the attacker
- The goal of the attack is to cause a denial of service for the users of the target system
- This is accomplished by prompting each of the members in the botnet to start sending messages to the same target at the same time
- This flood of incoming messages is intended to deplete the target resources to case the system to slow down or crash entirely, and to deny the service from its users.



https://ruggedtooling.com/what-are-ddos-attacks/

DDoS attack in blockchain

- In the blockchain space, the main DDoS threat is transaction flooding
- Most blockchains have a fixed capacity because they create blocks with a certain maximum size at regular intervals
- Anything that doesn't fit in the current block will be stored in mempools for consideration for the next block
- If an attacker sends many blockchain transactions containing negligible values to the network, they can fill up blocks with spam transactions causing legitimate transactions to sit in mempools

https://halborn.com/how-blockchain-ddos-attacks-work/

DDoS attack in blockchain

- The P2P nature of any blockchain system means these spam transactions sent multiple times by different nodes creating a congestion in the network
- On September 14, 2021, the Solana blockchain was offline for several hours
- The root cause of this issue was a DDoS attack caused by the launch of a new project on the blockchain
- When the project was launched, bots started generating large amounts of transactions that flooded the network

https://halborn.com/how-blockchain-ddos-attacks-work/

Question?