Blockchains after Bitcoin
– Big Data and IoT –

Winston Seah

Engineering and Computer Science

History of Blockchain

On 31 Oct 2008, *Satoshi Nakamoto* (real name?) proposed bitcoin

- Purely peer to peer electronic cash/digital asset transfer system
- Uses *blockchain* as the underlying technology

Bitcoin White Paper – https://bitcoin.org/en/bitcoin-paper

Bitcoin Network

- **B**itcoin network was launched in January 2009
- P2P electronic payment system that uses a cryptocurrency called ***bitcoin*** to transfer value over the internet or act as a store of value like gold and silver.
- Essentially, a bank run by an ad hoc network
  - Digital checks
  - Distributed transaction log

What is Blockchain?

Distributed Ledger Technology (DLT) – shared accounting system

- Distributed database – can insert & select data but cannot update / delete existing data
- Distributed processing – execute digital contracts
- Uses P2P technology, encryption and API

What is Blockchain?

Typically, a blockchain system is made up of:

- Transactions
- Immutable ledgers
- Decentralized peers
- Encryption processes
- Consensus mechanisms
- Optional Smart Contracts

Has mechanisms to make it hard to change historical records, or at least make it easy to detect when someone is trying to do so.

Blockchain Structure

- Nodes of the blockchain network have the same copy (duplicate) of the blockchain.
- Blocks are chained such that each block references the hash of its previous block.

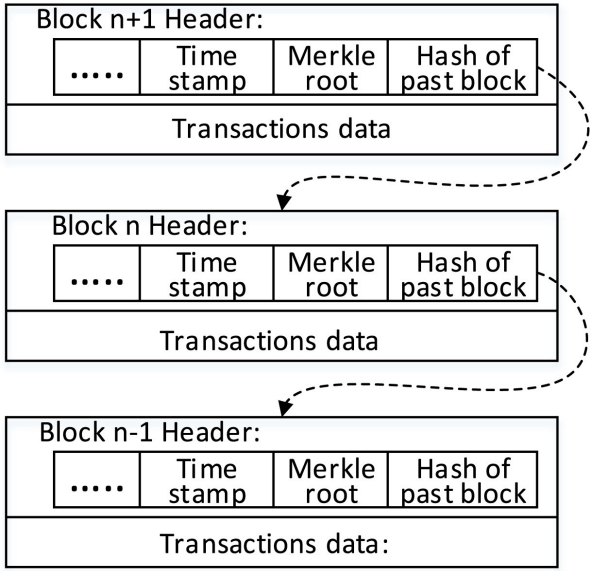
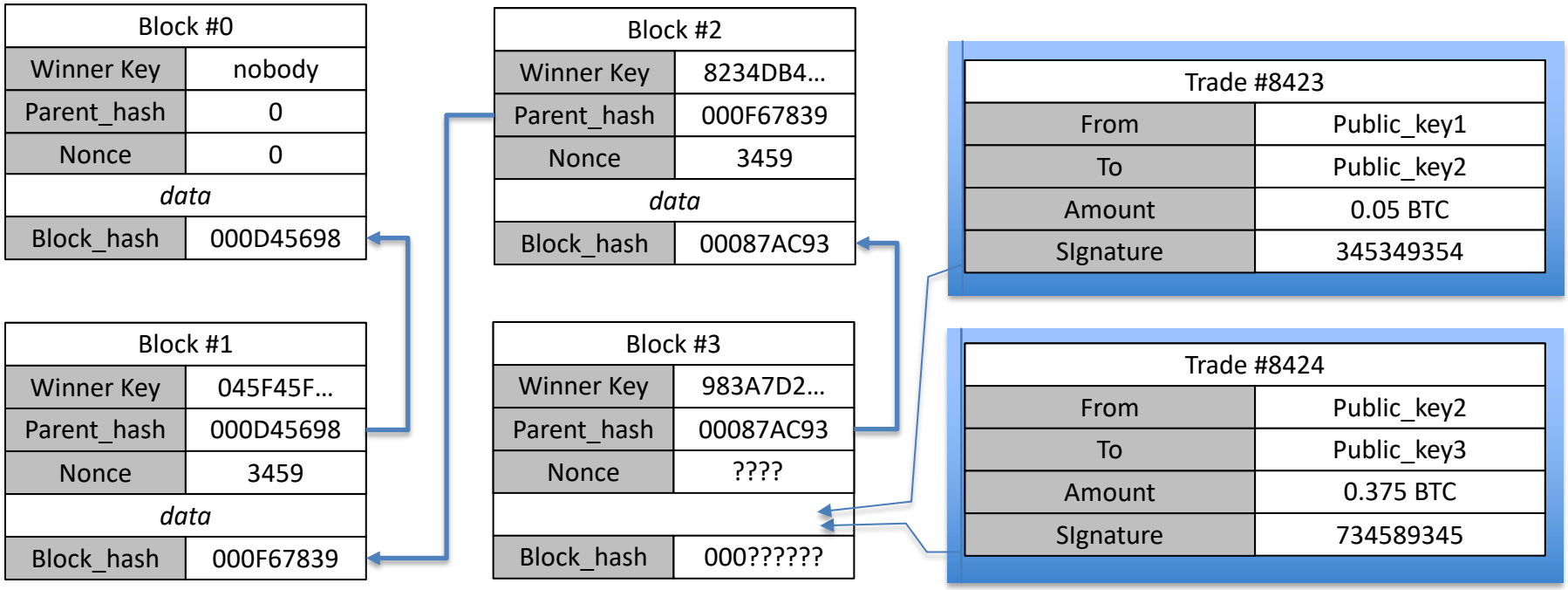


Figure source: Sanka *et al.*, "A survey of breakthrough in blockchain technology: Adoptions, applications, challenges and future research," *Computer Communications*, Vol 169, 2021, Pages 179-201.

Consensus Algorithms

- Used to create new blocks and add them to a blockchain.
- Types: Proof of Work (PoW), Proof of Stake (PoS), Proof of Elapsed Time (PoET), etc.
- PoW – most popular consensus protocol used by Bitcoin, Ethereum and many cryptocurrencies.

Adding new blocks - Bitcoin example



Bitcoin mining facility

Blockchain Types

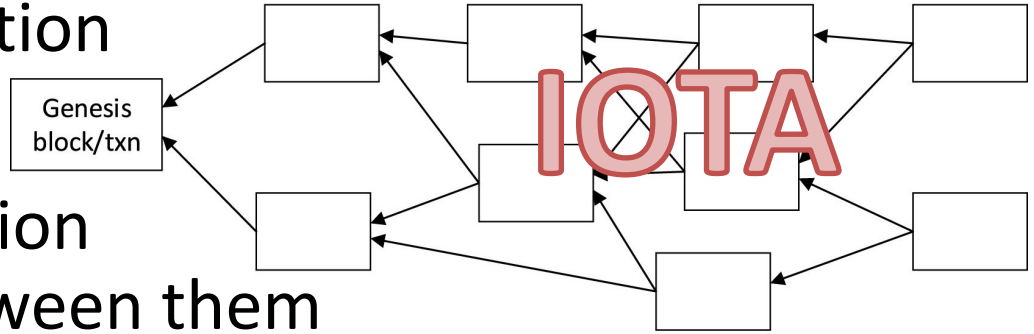
- Public
  - Open permissionless; anyone can participate
  - E.g., Bitcoin, Ethereum, most cryptocurrencies
- Private
  - Closed permissioned network; regulated control
  - Identified, trusted participants
  - E.g., Multichain and Blockstack

Blockchain Types

- Consortium or Federated
  - Group of organizations (consortium) to share data
  - Participants are known and require authorization to join the network
  - Assume little or no trust among its members
  - E.g., Hyperledger and Corda

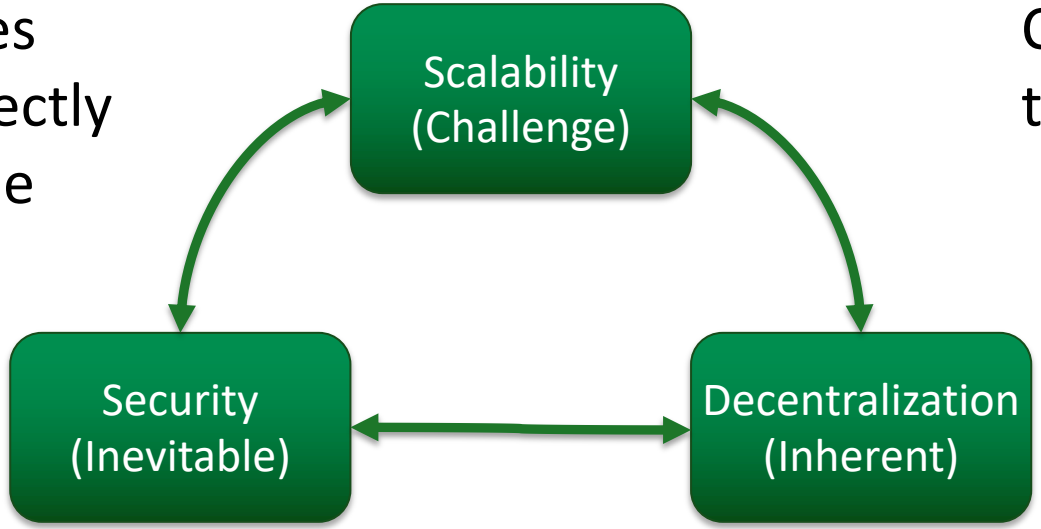
Blockchain Types

- Directed Acyclic Graph (DAG) based
  - Box → a transaction
  - Arrow → validation relationship between them
  - New transaction needs to validate its *parent* transaction and *parent-of-parent* transaction



Blockchain Scalability Trilemma

All 3 qualities cannot perfectly coexist at the same time



Choose-two tradeoff

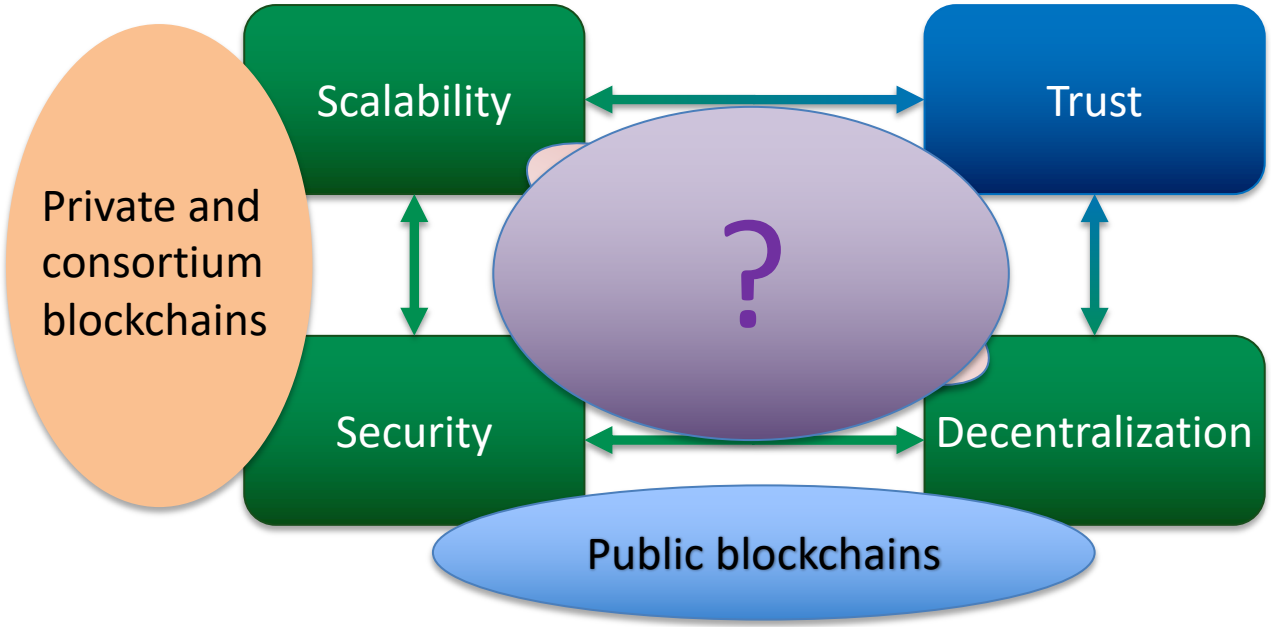
Blockchain Scalability Trilemma

- Trust is critical to blockchain scalability.
- Trusted parties

  → Less effort needed to validate transactions

  → Scalability can be achieved with less complex consensus algorithms, communications, and computations.

Blockchain Scalability Quadrilemma



Sharding

- Scaling method adapted from distributed database systems
- Partition blockchain network into groups called *shards*.

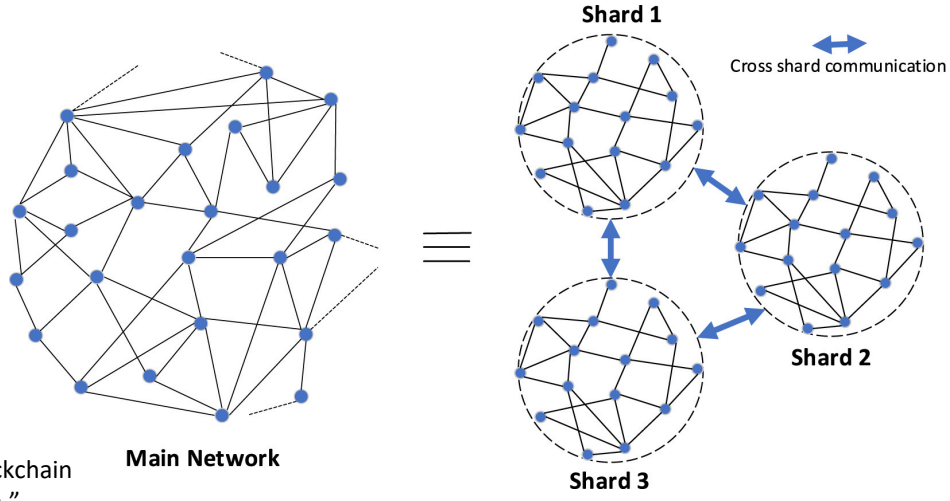


Figure source: Sanka and Cheung, "A systematic review of blockchain scalability: Issues, solutions, analysis and future research," *Journal of Network and Computer Applications*, Vol 195, 2021.

Sharding Advantages

- Each shard processes transactions and stores data in parallel.
- Allows parallel consensus and storage with increasing number of nodes.
- Reduces communication overheads in certain types of consensus networks.

Sharding Challenges

- Intra-consensus safety
  - Vulnerability to 1% attack
- Cross-shard atomicity
  - As number of shards increases, probability of cross-shard verification / transactions → 100%

Sharding Challenges

- General Improvements / Overheads, e.g.,
  - Transaction Latency can increase with added measures to deal with 1% attack
  - Inter-shard data transmission between miners / validators
  - Shards Ledger Management and Pruning to support cross-shard transactions

SideChains

- Secondary ledger (blockchain); attachment to main (primary) blockchain
- Allow asset transfer from main chain to sidechain at a predetermined rate for scalability

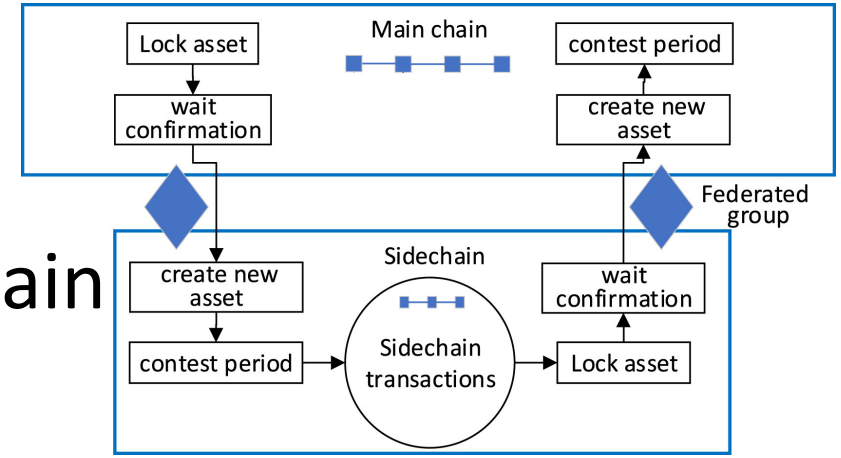


Figure source: Sanka and Cheung, "A systematic review of blockchain scalability: Issues, solutions, analysis and future research," *Journal of Network and Computer Applications*, Vol 195, 2021.

SideChains

- Low inter-dependency between main and side chain improves throughput, privacy, or security.
- Allows use of additional features unavailable on main chain, e.g., smart contract tokens.

X-chains & Off-chain Computation

- Cross-chains / X-chains
  - Similar to sidechains except they are pre-existing independent blockchains
- Off-chain Computation
  - offloads some tasks to relieve nodes in main chain from complex and time-consuming computation

What defines Big Data?

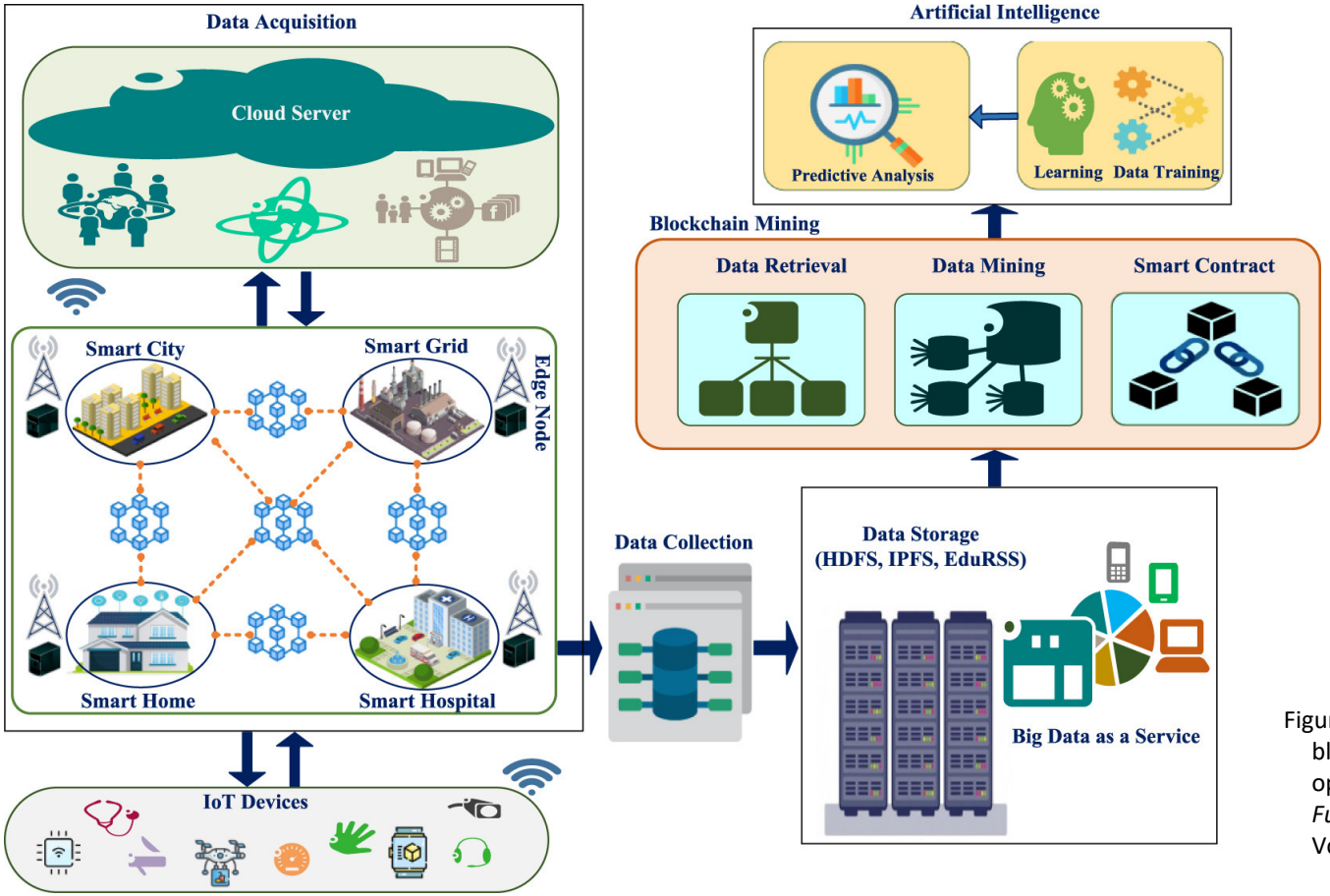
4Vs

- Volume – quantity of data
- Variety – types of data
- Velocity – generation speed of data
- Veracity – quality of data

Why Blockchain?

Big data contain information about us!!!

- Improve Security and Privacy
- Improve Data Integrity
- Prevent Fraud in financial sector
- Streamline Data Access
- Enhance Data Sharing
- Enhance Data Quality



Blockchain in Big Data Environment

Figure source: Deepa *et al.*, "A survey on blockchain for big data: Approaches, opportunities, and future directions," *Future Generation Computer Systems*, Vol 131, 2022, Pages 209-226,

Blockchain for IoT

- Enhanced Interoperability – processing and transforming IoT data for storage
- Improved Security – data (transactions) are encrypted and digitally signed
- Traceability and Reliability – data transactions are stored; immutability prevents tampering

Challenges in Blockchain for IoT

- Resource constraints of IoT devices
  - Decentralized consensus algorithms of blockchains often require extensive computing power and energy consumption
  - Bulky size of blockchain data
  - Blockchain originally designed for networks with good connectivity

Challenges in Blockchain for IoT

- Security Vulnerability
  - IoT systems themselves have poor security
  - Conventional encryption algorithms are too compute-intensive
  - Wireless connectivity open to attacks

Challenges in Blockchain for IoT

- Privacy leakage
  - IoT device identifiers, e.g. MAC, IP addresses, etc.
  - IoT device locations
- Incentive mechanisms
  - Who to pay for blockchain services?

Summary

- Blockchains have come a long way since bitcoin
- Inherent design makes it very attractive for many application scenarios, but …

Thank you!!!

Email:

winston.seah@ecs.vuw.ac.nz

Webpage:

https://www.ecs.vuw.ac.nz/~winston/