University of BATNA 2
Department of Computer Science

Chapter 3

Blockchain Architecture

Blockchain : Definition

- In the literature, several definitions have been proposed for the blockchain, here are two general definitions :

Blockchain at its core is a peer-to-peer distributed ledger that is cryptographically secure, append-only, immutable (extremely hard to change), and updateable only via consensus or agreement among peers. From a business point of view, a blockchain can be defined as a platform whereby peers can exchange values using transactions without the need for a centrally trusted arbitrator. A block is simply a selection of transactions bundled together in order to organize them logically. It is made up of transactions and its size is variable depending on the type and design of the blockchain in use. A reference to a previous block is also included in the block unless it's a genesis block. A genesis block is the first block in the blockchain that was hardcoded at the time the blockchain was started.

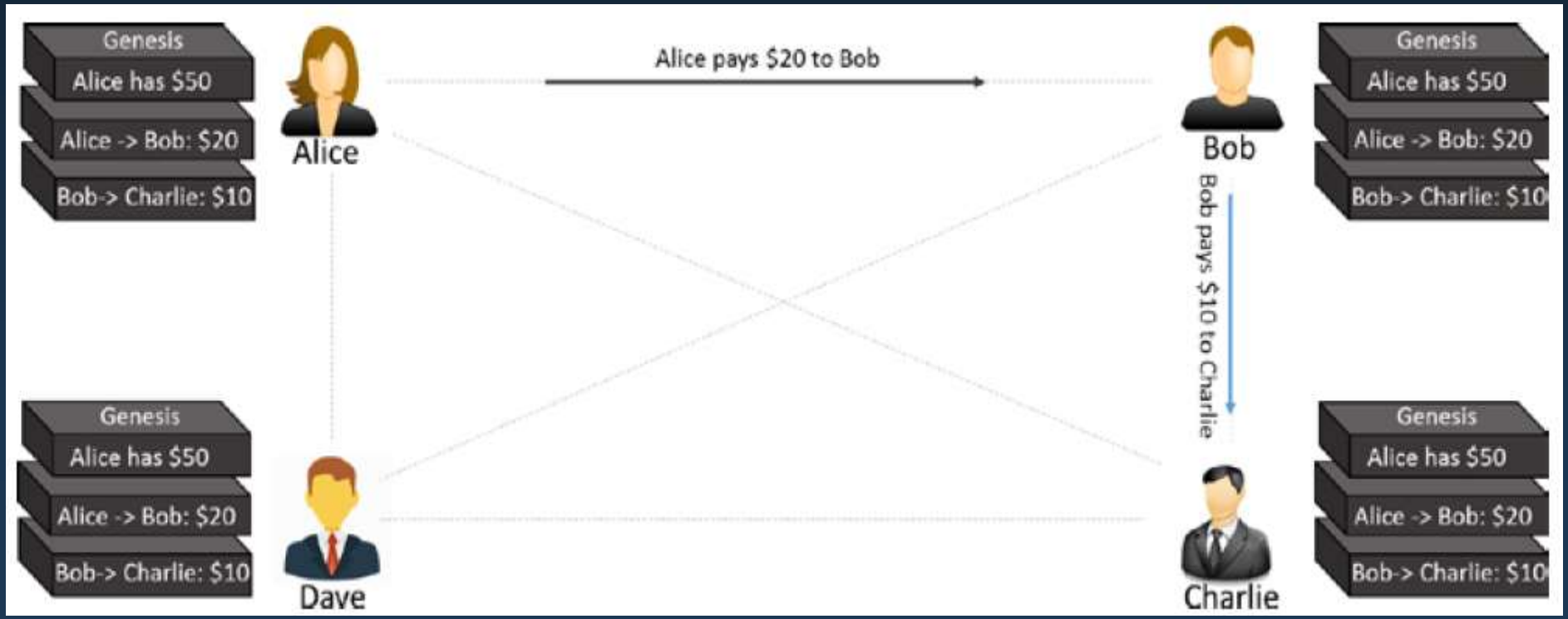
Blockchain : Definition

Blockchains are shared and distributed data structures or ledgers that can securely store digital transactions without using a central point of authority.The data structure is, in other words, a ledger that may contain digital transactions,data records and executables. Instead of managing the ledger by a single trusted center, each individual network member holds a copy of the records' chain and reach an agreement on the valid state of the ledger with consensus. The exact methodology of how consensus is reached is an ongoing area of research and might differ to suit a wide range of application domains. New transactions are linked to previous transactions by cryptography which makes blockchain networks resilient and secure. Every network user can check for themselves if transactions are valid, which provides transparency and trustable, tamper-proof records.

- These definitions allow to bring out the terminology and the set of concepts constituting the basic bricks of the blockchain technology.

Transaction

- Transactions (المعاملات) are asset transfers in a blockchain sytem. Depending on the system, transactions may contain any kind of assets such as financial value (Debit, Credit, Transfer) , health data, log record, identity information, voting data, program instruction, ….



Taxonomy of Blockchain

- Public : No access restriction, Permessionless, Anyone can act as a node, Open, Perform transaction, Participate in the consensus (الإجماع) process, Decentralised, Transparent, No intervention of a central authority, ...

  Example : Ethereum, Bitcoin

- Private : Access restriction (only invited), Permessionned, Partially decentralised, Not transparent, Intervention of a central authority, Generally single oraganisation, …

  Example : Ripple

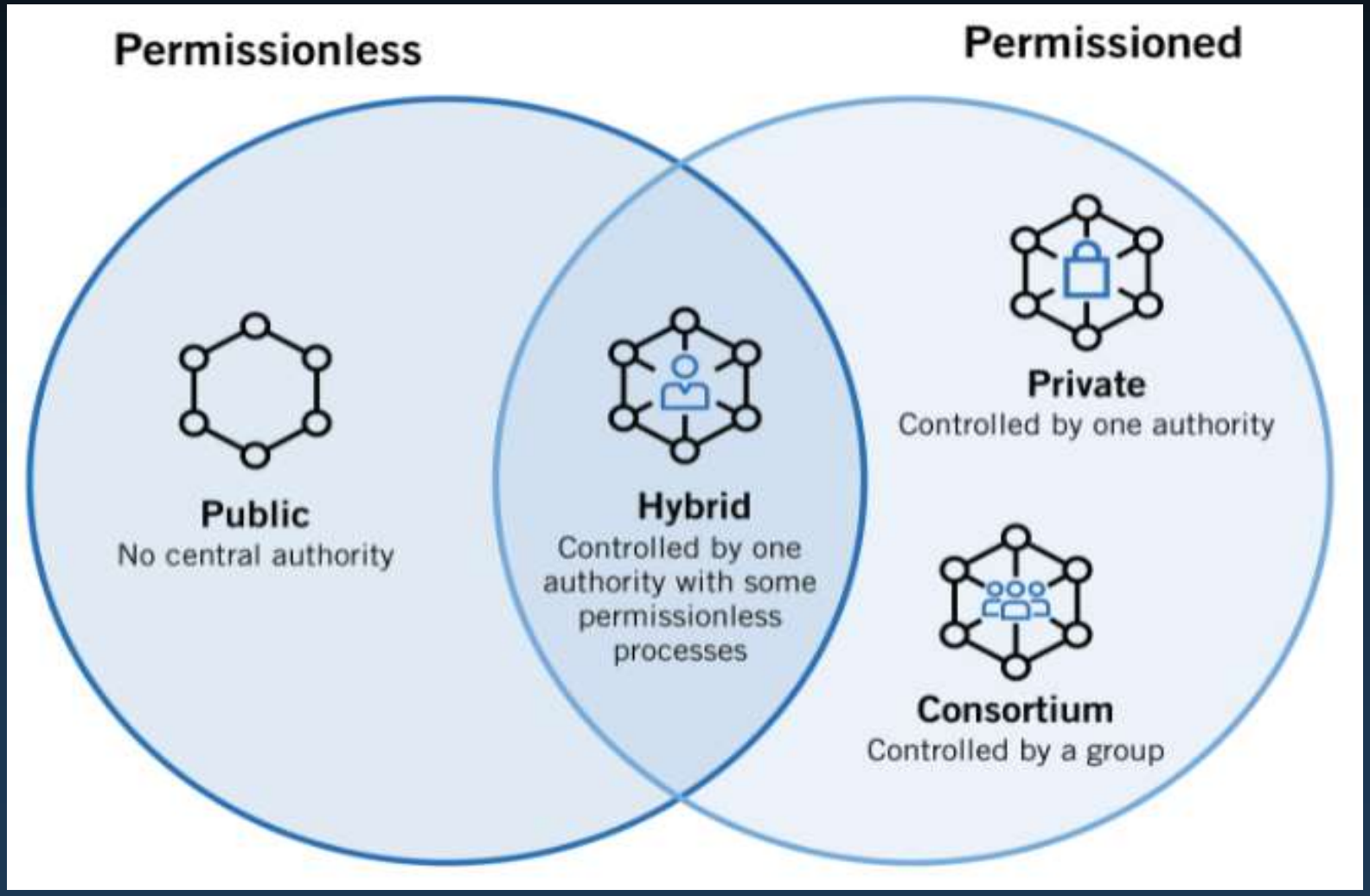
- Consortium (Federated) : A combination of both public and private blockchains. It's appropriate where multiple organizations work together.

  Example : Quorum

- Hybrid : It's controlled by a single organization, but with a level of oversight performed by the public blockchain, which is required to perform certain transaction validations.

  Example : IBM Food Trust

Taxonomy of Blockchain



Blockchain principle

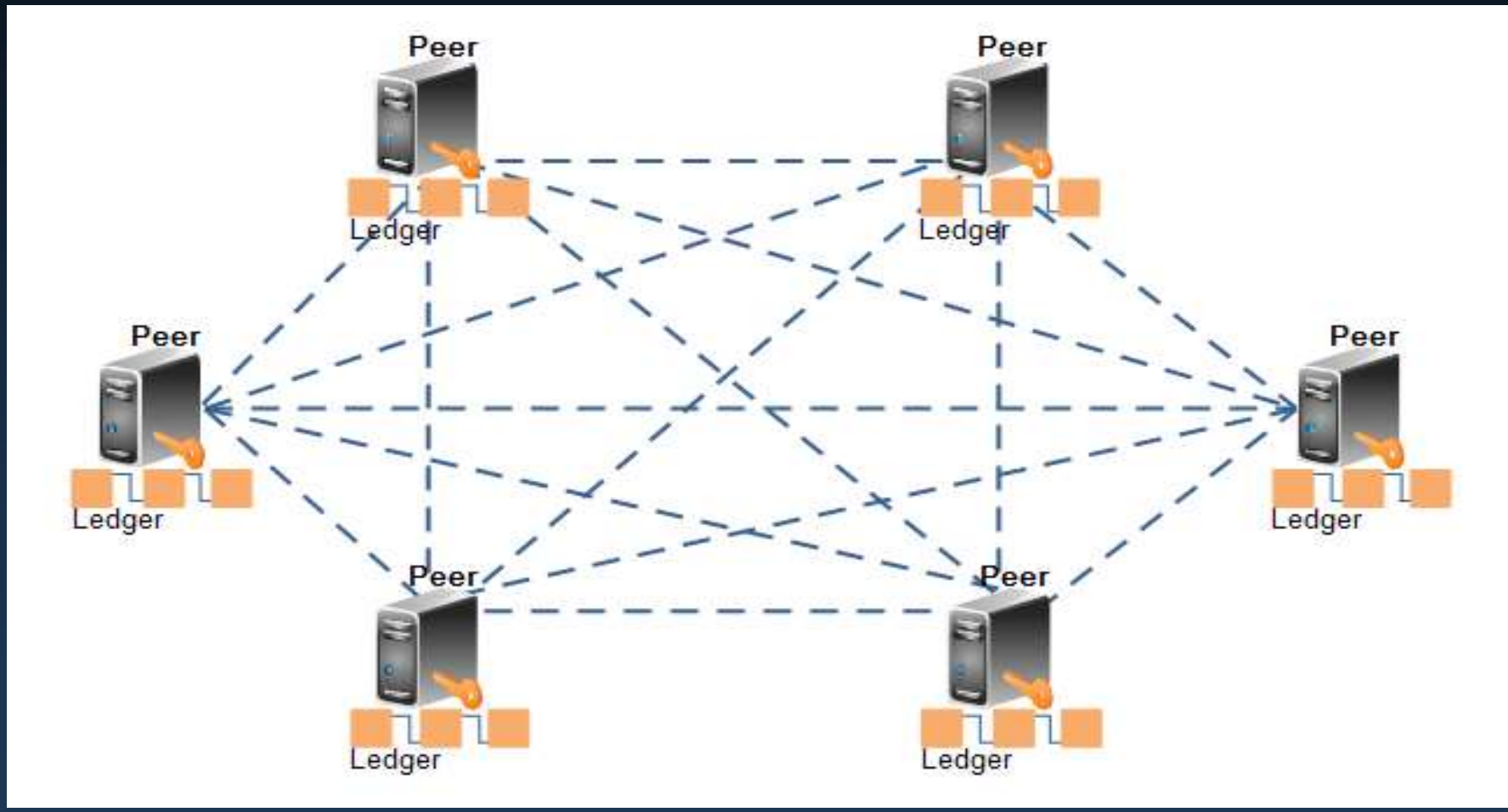
- According to Satoshi Nakamoto, the following steps describe the general principle of the blockchain

1. New transactions are created
2. The new transactions are broadcast to all nodes in the P2P network.
3. Each node collects new transactions into a block.
4. Following the same consensus protocol, each node works on finding a valid block : Mining
5. The miner who find a valid block will be rewarded for its work. The reward is expressed in a specific currency (Cryptomoney) : Ether, Bitcoin, Wei, ...
6. When a node finds the proof-of-work, it broadcasts the block together with the solution to all nodes.
7. Nodes accept the block only if all transactions in it are valid and if the coins were not already spent.
8. Nodes express their acceptance of the block by creating the next block in the chain and using the hash of the accepted block as the previous hash.

Blockchain Network : Peer-To-Peer

- The blockchain at its core is a peer-to-peer distributed ledger. Instead of managing the ledger by a single trusted element, each individual network node manages and holds a copy of the block's chain.
- Three kinds of nodes can be distinguished :
  - Full node : Full node maintains a complete copy of the blockchain, able to verify all transactions and verifies a newly created block and adds it to the blockchain.
  - Mining node : Mining node must be a full node and ensure the mining process.
  - Lightweight node : A node that does not store or maintain a copy of the blockchain, must pass their transactions to full nodes and can verify if a transaction is present and valid in a block.These nodes are generally used for development, the wallet (حافضة) is an example of light node.

Blockchain Network : Peer-To-Peer



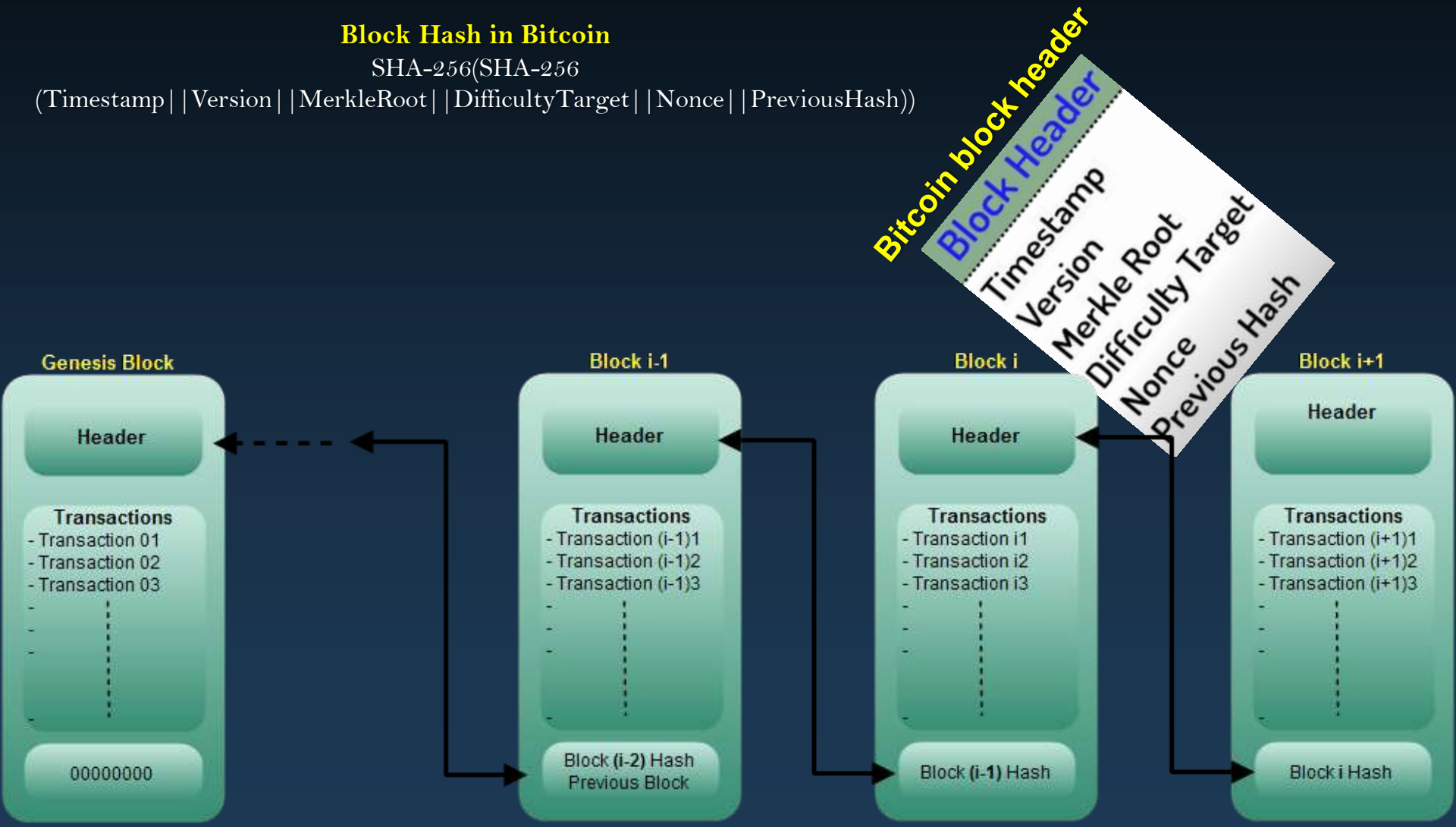
Blockchain Structure

- As a data structure, the blockchain is a linked list of elements. The element, composed of several fields, is the block.
- The blocks are not linked throw addresses as in data structure, but the link value is provided by a hash function.
- In addition to the descriptive fields that constitute the header, the block body contains a set of structured data : Transaction.
- Every block header includes hash value of the previous block. So, each block is linked to the previous, constituting a chain of blocks.
- When data in a block changes, the block hash also changes. consequently, the next block's hash output also change in the chain.

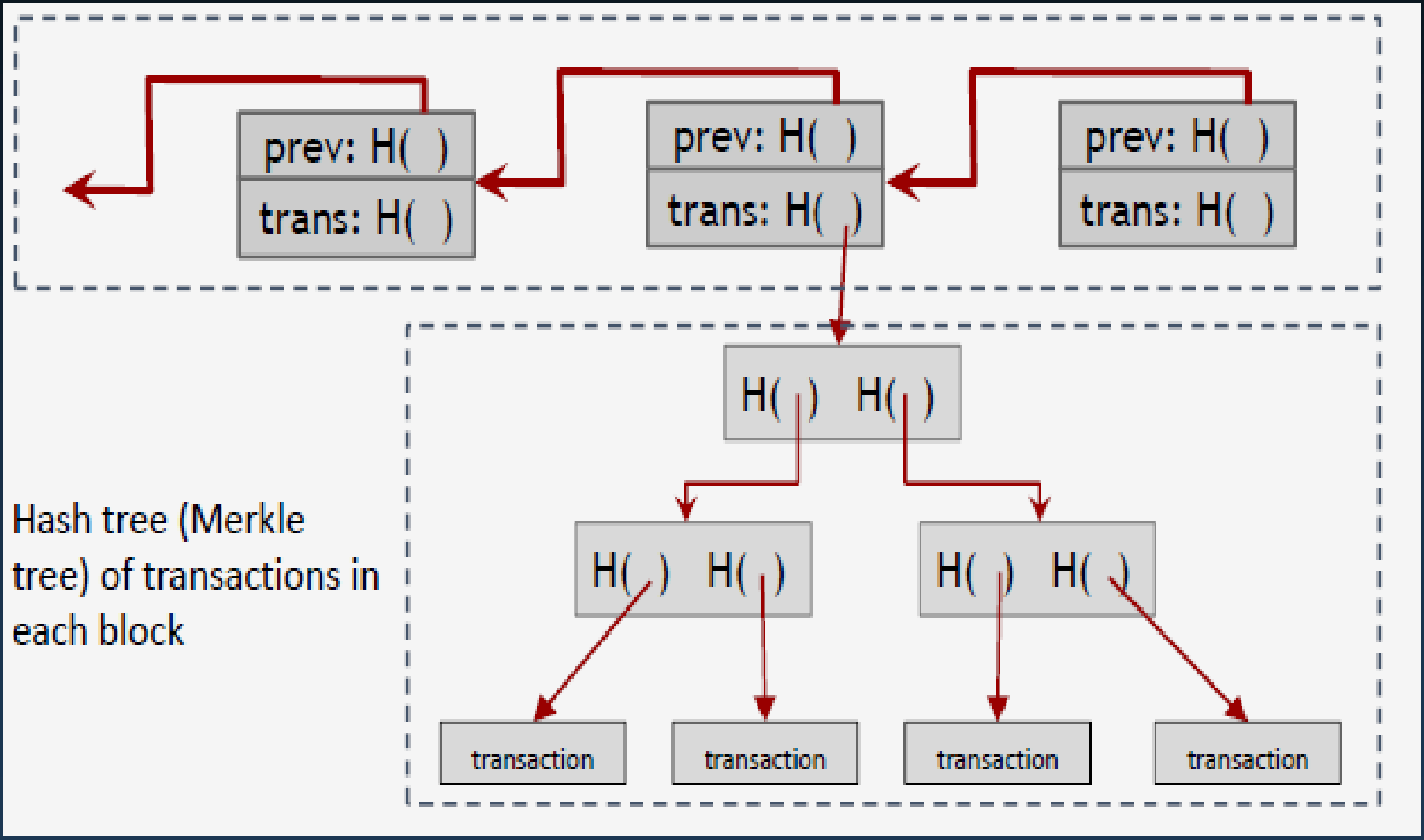
Blockchain Structure

Block Hash in Bitcoin

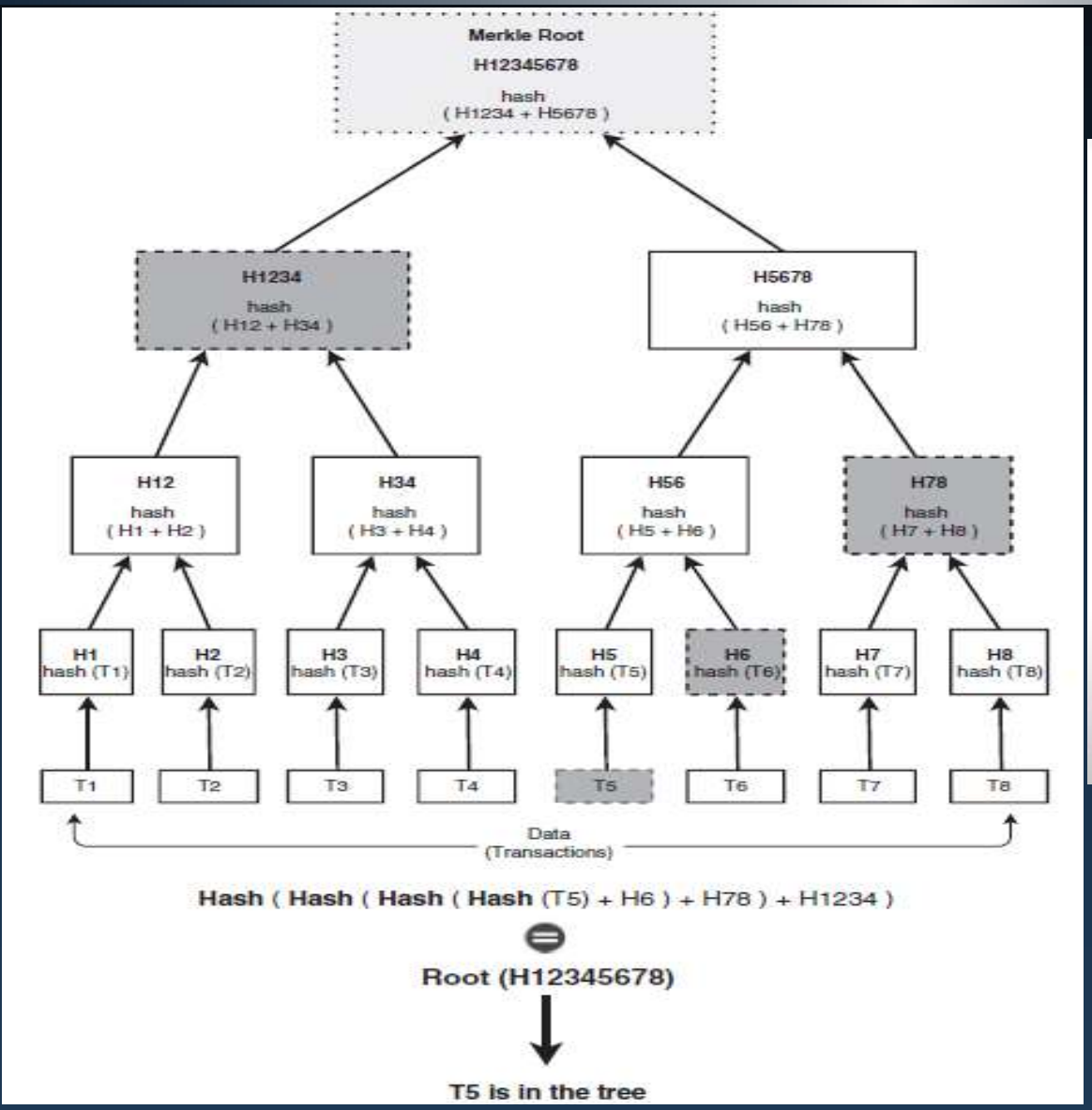
SHA-256(SHA-256
(Timestamp||Version||MerkleRoot||DifficultyTarget||Nonce||PreviousHash))

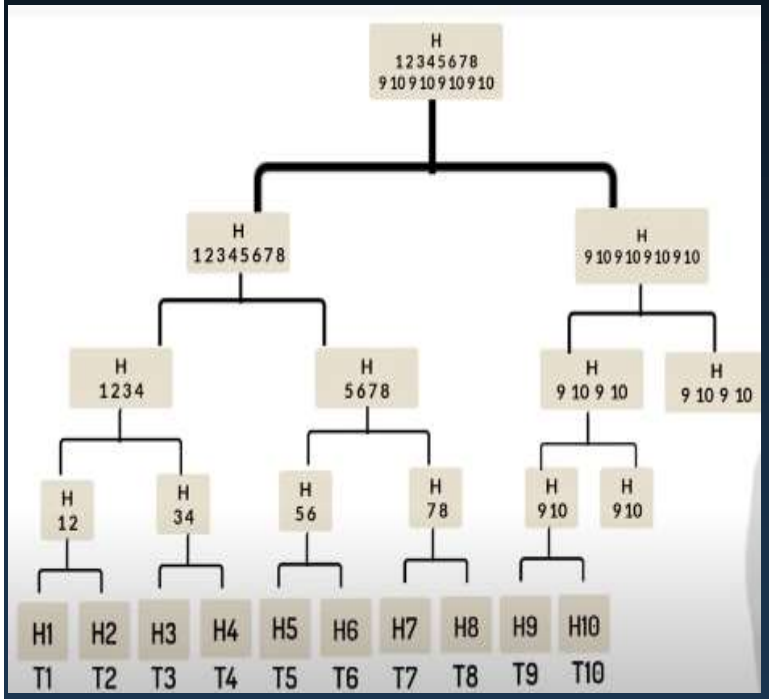


Blockchain Structure : *Merkle Root*



Blockchain Structure : *Merkle Root*





Consensus Algorithms

- The consensus algorithm in Blockchain is used to solve the problem of ensuring *data consistency* in the presence of several failure nodes in a distributed system.
- Data consistency : Ensuring the insertion of valid block, so valid transactions.
- The famous Byzantine Generals Problem is raised in blockchain : *Consensus Algorithms*

Byzantine generals problem

The Byzantine generals problem can be described as follows :

- ✓ *Byzantine is the capital of the ancient eastern Roman Empire.*
- ✓ *There are several fiefs in Byzantine, each stationed by a general and his soldiers to resist foreign enemies.*
- ✓ *When facing enemies, each general can give two orders: attack or retreat.*
- ✓ *Only when all honest generals agree on an order to attack or retreat can they minimize casualties and win a war.*
- ✓ *However, Byzantine is so vast that these generals cannot discuss the order together because they have to guard in their fiefs separately.*
- ✓ *Therefore, the orders from generals are delivered via signalmen.*
- ✓ *The generals eventually make their final decisions (attack or retreat) by sending their orders to the other generals and collect the orders from the other generals.*
- ✓ *In this scenario, we assume that the signalmen are honest.*
- ✓ *However, some of these generals are traitors, who may send wrong orders or send different orders to different generals and ultimately undermine the overall decision of the honest generals.*

In summary, the Byzantine generals problem can be formulated as a problem of getting honest generals to reach a consensus in the presence of several traitors.

Blockchain Consensus Algorithms

- Consensus in a blockchain can be achieved by applying various proof-based algorithm
- Consistency and liveness are the fundamental issues to be considered when designing a Blockchain consensus algorithm
- Earliest deployed consensus algorithms : Proof of work (PoW), proof of stake (PoS), and practical Byzantine fault tolerance (PBFT).
- Some emerged consensus algorithms :
  - ✓ Variants of the original consensus algorithms, e.g., Bitcoin-NG and Algorand , which are the improvements of PoW and PBFT, respectively;
  - ✓ Combinations of the original consensus algorithms, e.g., delegated BFT (DBFT), which is the combination of PoS and PBFT
  - ✓ Directed Acyclic Graph DAG-based consensus algorithms, e.g., Byteball and Hashgraph

Proof of work : PoW

- PoW is the most popular and widely used consensus algorithm, which is used in Bitcoin.
- Under PoW the participant is called Miner, wich peforms a hard work called Mining (التعدين).
- The miners perform a computational task, i.e., solve a cryptographic puzzle problem in order to obtain the right to generate a new block.
- The nodes perform puzzle solving by finding a specific hash function.
- Each miner tries to solve the hash value, i.e., they try to figure out a particular value as a nonce to meet a predefined hash condition, e.g., finding the nonce value that will make the first 30 bits of its hash to zero, i.e., finding hash values that are smaller than or equal to a certain target value.
- Whenever a miner finds the target hash value, it broadcasts the current block to the whole network.
- All other nodes verify the correctness of its hash value. If the block is legit, all nodes append that new validated block to their blockchain.
- The miner receive a reward for completing the hard computational task.

Proof of work : PoW

- Example : The task is to find a hash value meeting the following target criteria (known as the difficulty level) :

```
SHA256("blockchain" + Nonce) = Hash Digest starting with "000000"
```

- In this example, the text string "blockchain" is appended with a nonce value and then the hash digest is calculated

```
SHA256("blockchain0") =
0xbd4824d8ee63fc82392a6441444166d22ed84eaa6dab11d4923075975acab938
(not solved)

SHA256("blockchain1") =
0xdb0b9c1cb5e9c680dfff7482f1a8efad0e786f41b6b89a758fb26d9e223e0a10
(not solved)

…

SHA256("blockchain10730895") =
0x000000ca1415e0bec568f6f605fcc83d18cac7a4e6c219a957c10c6879d67587
(solved)
```

- To solve this puzzle, it took 10,730,896 guesses (completed in 54 seconds on relatively old hardware, starting at 0 and testing one value at a time).

Proof of work : PoW

- In this example, each additional “leading zero” value increases the difficulty. By increasing the target by one additional leading zero (“0000000”), the same hardware took 934,224,175 guesses to solve the puzzle (completed in 1 hour, 18 minutes, 12 seconds) :

```
SHA256("blockchain934224174") =
0x0000000e2ae7e4240df80692b7e586ea7a977eacbd031819d0e603257edb3a81
```

Proof of work : PoW with SHA256

Bitcoin uses

$$\begin{array}{rcl} \{0,1\}^{2^{64}} & \to & \{0,1\}^{256} \\ x & \mapsto & \mathrm{SHA256}(\mathrm{SHA256}(x)) \in \{0,1\}^{256} \end{array}$$

- given a "target" $T \in [0, 2^{256})$
- to "mine" block-data:

  UNTIL hash $< T$

  nonce = next nonce

  hash = H(block-data || nonce)

No better way than guessing. Probability of success for one nonce: $T/2^{256}$

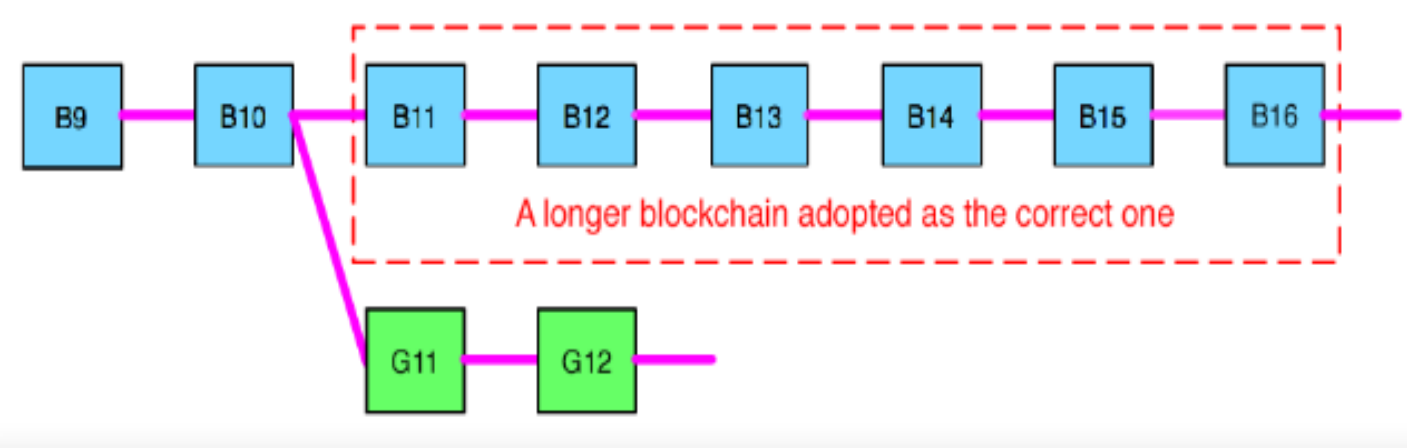
T is readjusted every 2016 blocks, to keep producing a block every 10 min

$$\texttt{difficulty} \leftarrow \texttt{difficulty} \cdot \frac{\texttt{2 weeks}}{\texttt{time to mine last 2016 blocks}}$$

difficulty is proportional to 1/T

Proof of work : Blockchain branches

- It is possible that multiple blocks are validated and will be published at approximately the same time
  - ↳ The generation of branches
- In PoW protocol, the chain is authentic if it becomes long thereafter. Consider two forks created by simultaneously validated blocks B11 and G11. The mining is performed by miners until a longer chain is found. The longer chain is formed by B11, B12, B13, …. So, the miners on G11 switch to the longer chain



Proof of work : Block validation

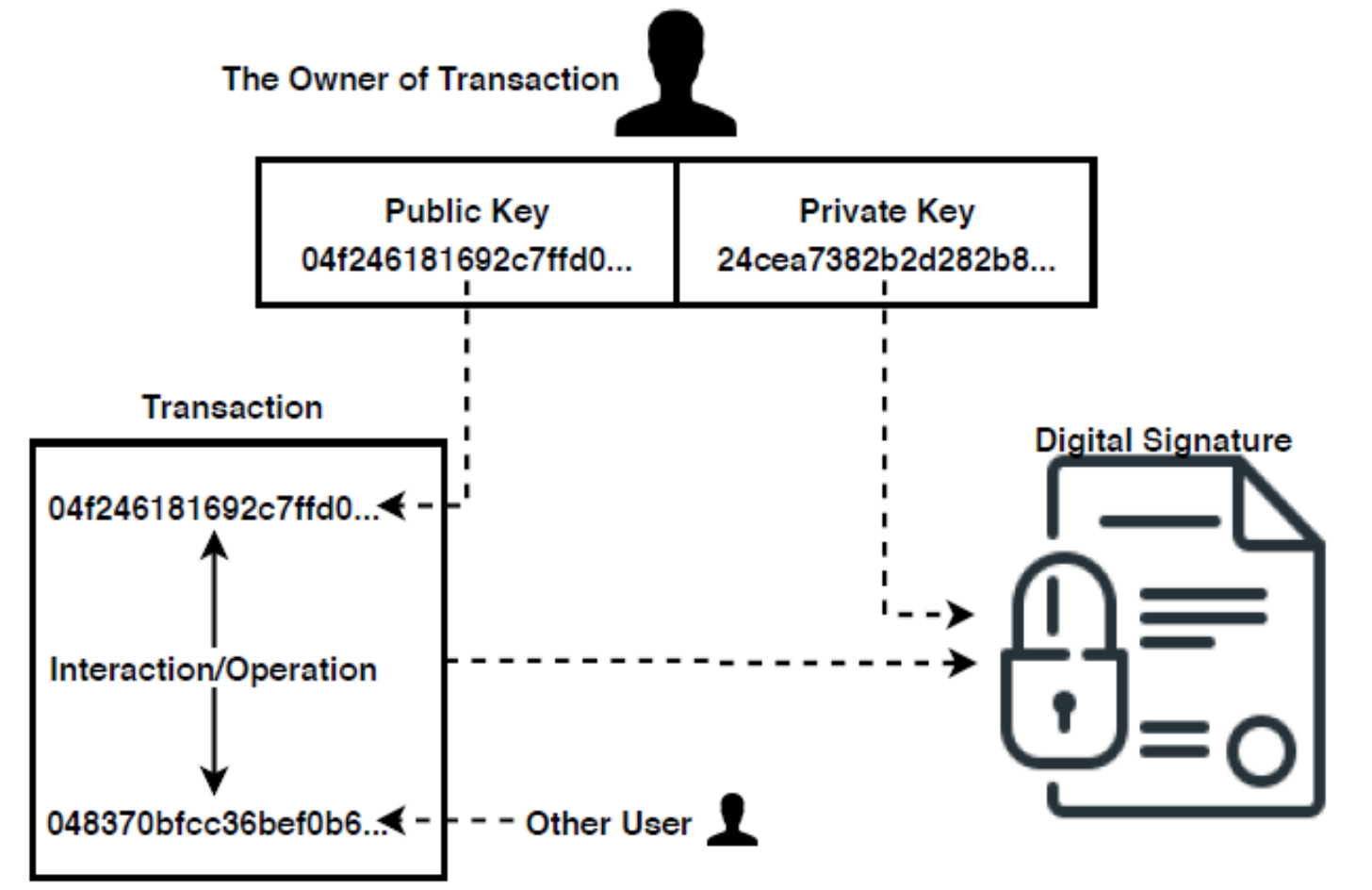
1. Check syntactic correctness
2. Make sure neither in or out lists are empty
3. Size in bytes <= MAX_BLOCK_SIZE
4. Each output value, as well as the total, must be in legal money range
5. Make sure none of the inputs have hash=0, n=-1 (*coinbase* transactions)
6. Check that nLockTime <= INT_MAX[1], size in bytes >= 100[2], and sig opcount <= 2[3]
7. Reject "nonstandard" transactions: scriptSig doing anything other than pushing numbers on the stack, or scriptPubkey not matching the two usual forms[4]
8. Reject if we already have matching tx in the pool, or in a block in the main branch
9. For each input, if the referenced output exists in any other tx in the pool, reject this transaction.[5]
10. For each input, look in the main branch and the transaction pool to find the referenced output transaction. If the output transaction is missing for any input, this will be an orphan transaction. Add to the orphan transactions, if a matching transaction is not in there already.
11. For each input, if the referenced output transaction is coinbase (i.e. only 1 input, with hash=0, n=-1), it must have at least COINBASE_MATURITY (100) confirmations; else reject this transaction
12. For each input, if the referenced output does not exist (e.g. never existed or has already been spent), reject this transaction[6]
13. Using the referenced output transactions to get input values, check that each input value, as well as the sum, are in legal money range
14. Reject if the sum of input values < sum of output values
15. Reject if transaction fee (defined as sum of input values minus sum of output values) would be too low to get into an empty block
16. Verify the scriptPubKey accepts for each input; reject if any are bad
17. Add to transaction pool[7]
18. "Add to wallet if mine"
19. Relay transaction to peers
20. For each orphan transaction that uses this one as one of its inputs, run all these steps (including this one) recursively on that orphan

https://en.bitcoin.it/wiki/Protocol_rules

Blockchain Public Key & Private Key & Address

- The concept of ownership on a blockchain system is primarily comprised of three interconnected elements :
  - Digital keys (Public key and Private key)
  - Cryptocurrency addresses
  - Digital signatures (ECDSA)
- Transactions are basic units or atomic events of blockchain protocols. Blockchain protocols usually has their own type of assets, which are transferred through transactions.
- The figure shows a basic transaction between two users is shown. the transaction is between 04f2461816 92c7d0... and 48370bfcc36bef0b6... addresses. Each address represents a real world user without revealing any personal information.

Blockchain Public Key & Private Key & Address



Blockchain Public Key & Private Key & Address

Digital signature scheme. A digital signature scheme consists of the following three algorithms:

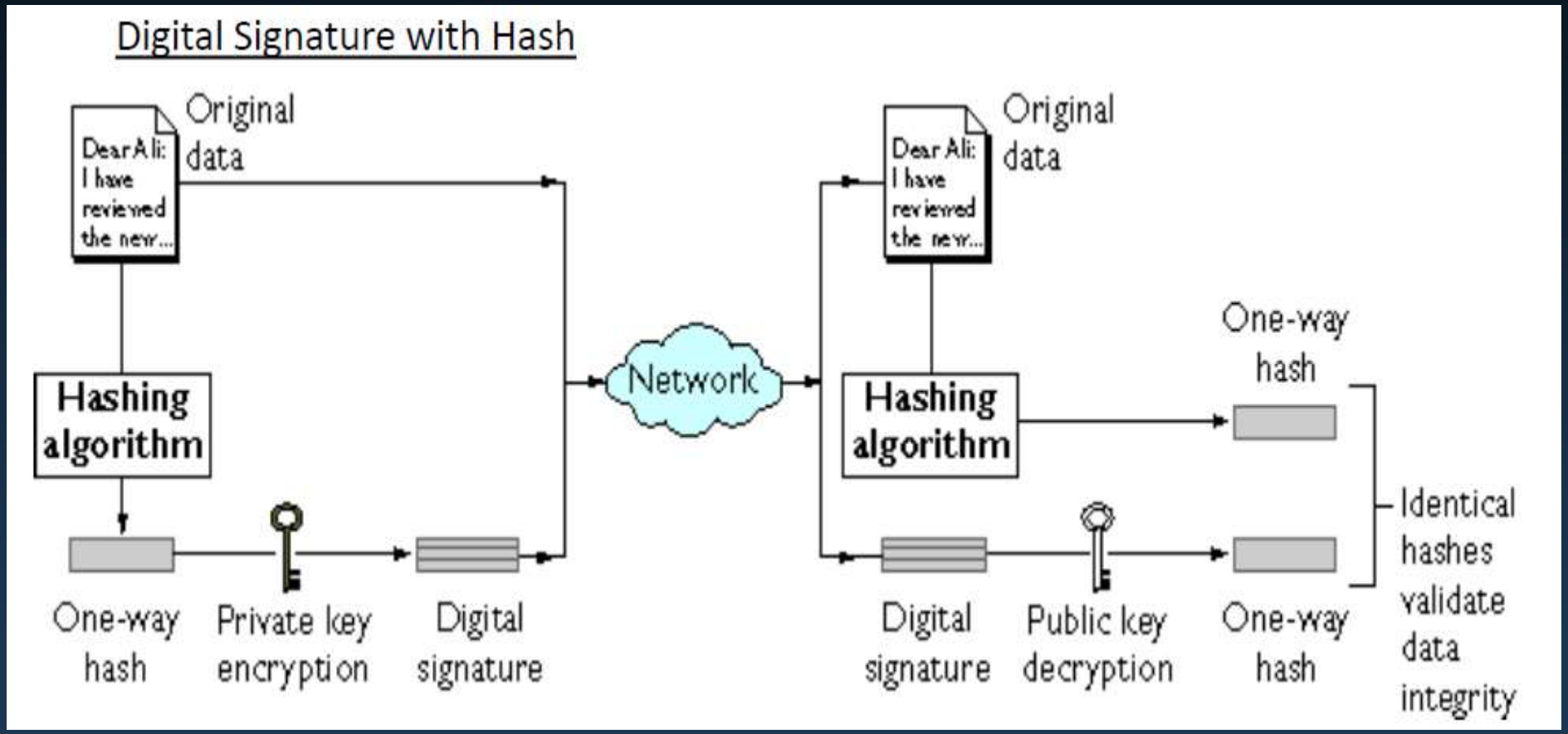
- **(sk, pk) := generateKeys(*keysize*)** The generateKeys method takes a key size and generates a key pair. The secret key *sk* is kept privately and used to sign messages. *pk* is the public verification key that you give to everybody. Anyone with this key can verify your signature.
- **sig := sign(*sk, message*)** The sign method takes a message and a secret key, *sk*, as input and outputs a signature for *message* under *sk*
- **isValid := verify(*pk, message, sig*)** The verify method takes a message, a signature, and a public key as input. It returns a boolean value, *isValid*, that will be ***true*** if *sig* is a valid signature for *message* under public key *pk*, and ***false*** otherwise.

We require that the following two properties hold:

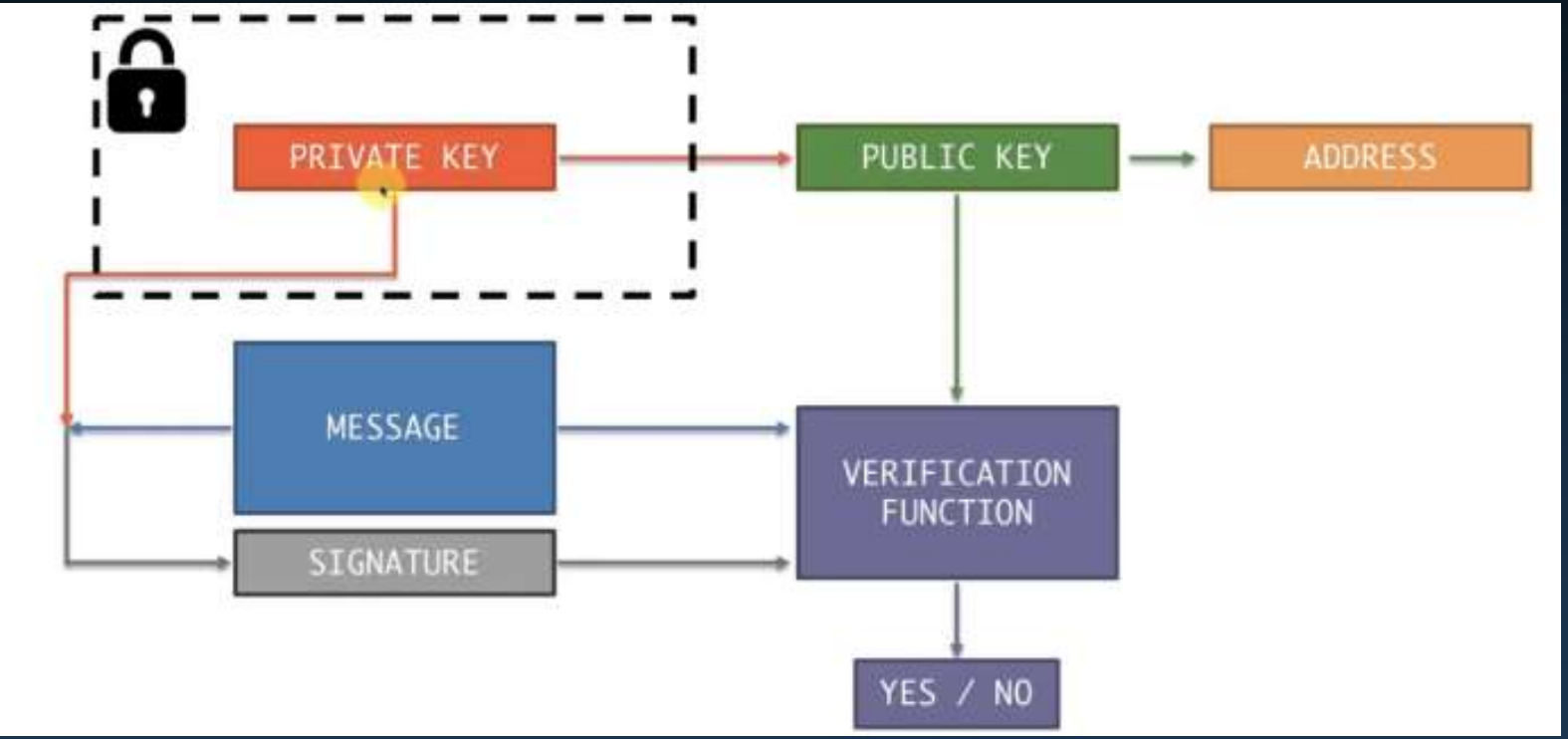
- *Valid signatures must verify*

  verify(*pk*, *message*, **sign**(*sk*, *message*)) **== true**
- Signatures are ***existentially unforgeable***

Blockchain Public Key & Private Key & Address



Blockchain Public Key & Private Key & Address



Blockchain Public Key & Private Key & Address