APS

5. Block chain Data structure.

5.1 Blockchain Architecture:

INTRODUCTION:

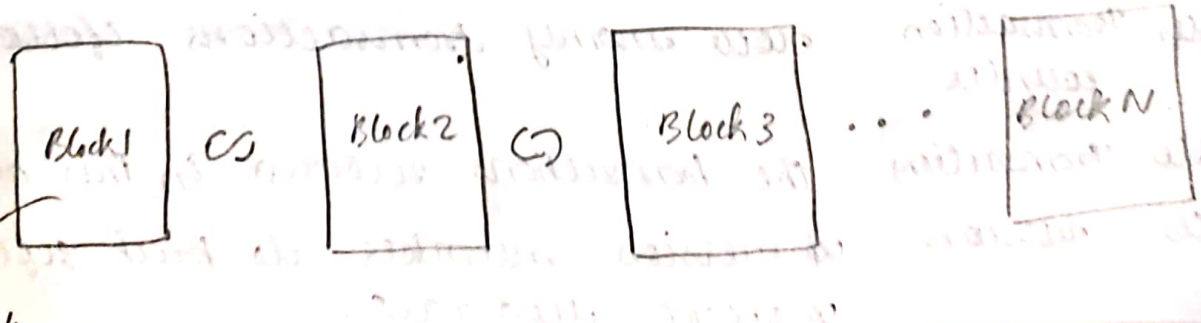
Blockchain :-

→ The term blockchain was first described back in 1991.

→ A group of researchers went to timestamp digital documents so that they could not be back-dated or changed.

→ The technique was adapted & reinvented by Satoshi Nakamoto. In 2008, Nakamoto created the first cryptocurrency, the block-chain based project called Bitcoin.

→ A Blockchain is a chain of blocks which contains information. The data which is stored inside a block depends on the type of blockchain.



→ Nevertheless, blockchain is a combination of computers linked to each other instead of a central server, meaning that the whole network is decentralised.

→ each block can be recognised by a hash, created by utilising the SHA256 cryptographic hash algorithm on the header of the block.

→ each block mentions a former block, also identified as the parent block, in the 'previous block hash' field, in the block header.

→ Block:

- is a package data structure.
- is composed of a header which includes metadata, accompanied by a lengthy record of transactions that advance its size.

Size	Field	Description
4 bytes	Block size	The size of the block, in bytes.
80 bytes	Block header	several fields from the block header
1-9 bytes	Transaction counter	how many transactions follow
variable	Transactions	The transactions recorded in this block.
4 bytes	version	A version number to track software / protocol upgrades.

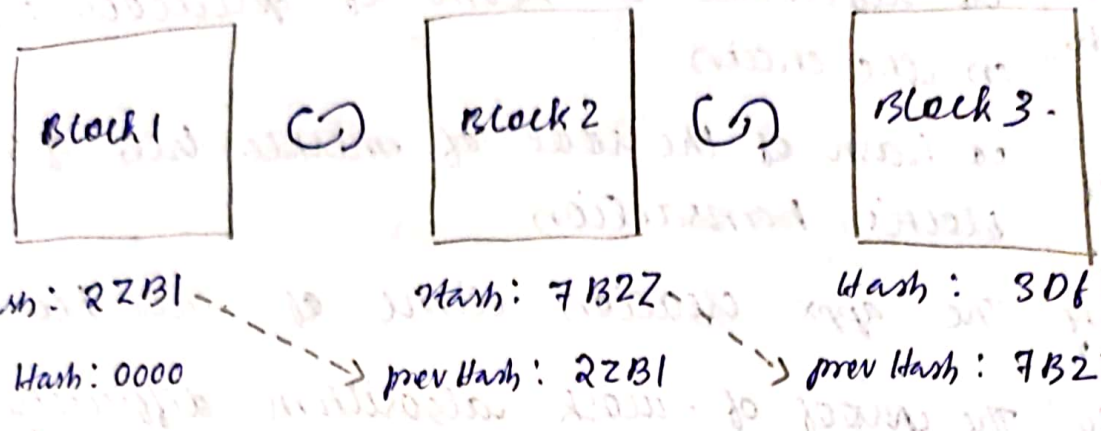
32 bytes	Previous block hash	A reference to hash of previous block in the chain.
32 bytes	Merkle root	A hash of the root of merkle tree of this block's transaction
4 bytes	Timestamp	The approx creation time of this block.
4 bytes	Difficulty target	The proof-of-work algorithm difficulty target for this block.
4 bytes	Nonce	A counter used for the proof-of-work algorithm.

→ Primitive identifier of a block is its cryptographic hash. It is also known as a digital fingerprint which is built by hashing the block header twice by through SHA256 algorithm. The resulting 32-byte hash is described as block hash.

→ Genesis block : ⊖ first block in blockchain.
- ⊖ built in year 2009
- ⊖ universal parent of all blocks in blockchain.

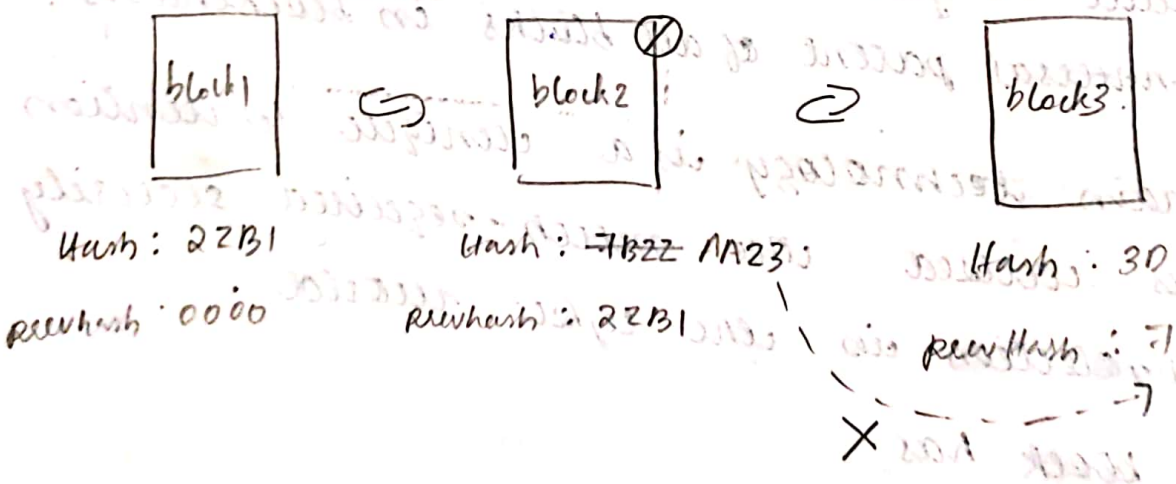
→ Blockchain technology is a unique invention that has caused the much-required security and protection in the cyber world.

→ each block has
1. Data
2. Hash
3. Hash of previous block.



→ Hence all block contains names of previous blocks. This is the technique that makes a blockchain so secure.

Assume an attacker is able to change the data present in block 2. Correspondingly, the hash of the block also changes. But block 3 still contains the old Hash of block 2. This makes block 3 and all succeeding blocks invalid as they do not have correct hash of previous block.



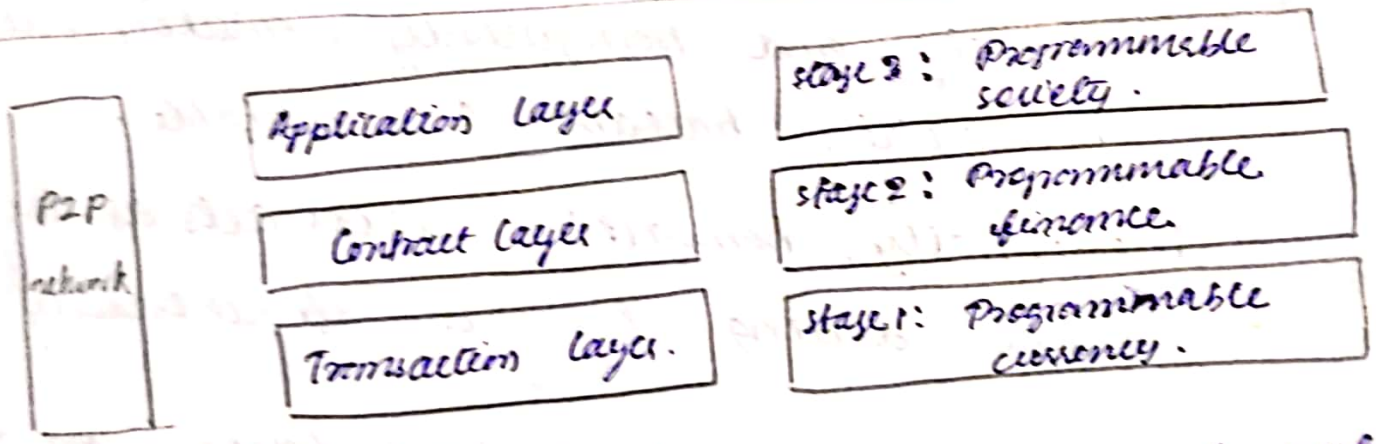
Blockchain Architecture:

→ Blockchain is the underlying technology of bitcoin.

→ In 2008, a scholar named "satoshi Nakamoto" proposed a digital currency called Bitcoin.

→ in the absence of any authoritative intermediates, people who donot trust can pay directly in Bitcoin.

→ Blockchain technology is a new application made of computer technology, such as distributed data storage, point-to point transmission, & consensus mechanism and encryption algorithm.



→ This is the blockchain framework, which contains B.C. architecture layers.

→ The 3-tier framework not only represents 3 types of BC data, but also 3 stages of BC development.

→ 1. The bottom layer is transaction layer, corresponding to BC development stage 1.0

- it is the foundational layer where all transactions are created, verified and processed.
- (i) when a user initiates a transaction, it is first created and then broadcasted to the network for verification.
- (ii) Nodes in the network validate the transaction, confirming that the sender has sufficient funds.
- (iii) once verified, the transaction is grouped with others to form a block, which is then added to the blockchain.
- this layer ensures data integrity, security and transparency, making every transaction traceable & immutable.
- essentially, transaction layer acts as the engine driving BC's core functionality.

→ 2. The middle layer is the contract layer, corr. to 2.0 phase of BC development.

- where, smart contracts are managed and executed.
- smart contracts are self-executing

programs that automatically carry out specific actions when certain conditions are met.

- This layer helps automate processes, removing need for intermediaries like banks or lawyers, which saves time & reduces costs.
- The contract layer plays a key role in enabling complex applications like decentralized finance and automating tasks.

→ 3. The top layer is the application layer, or to 3.0 phase of BC development.

- where, users interacts with the system.
- provides interface & tools needed to use BC based applications, such as wallets, decentralised apps and web platforms.
- This layer translates complex BC func. to simple actions that users can perform, like sending cryptocurrency, accessing smart contracts...
- designed to make BC technology accessible and user friendly.
- Essentially, Appli layer is what brings BC technology to life for end-users.

6) Core components of BC Architecture :-

These are core BC arch. components :

(i) Node :-

- individual computers / devices that participate in BC network.
- Types :

  Full Nodes – maintain complete copy of the BC

  Light Nodes – store only part of BC.

(ii) Ledger :-

- records all transactions across the network.
- every node has copy of this ledger, ensuring transparency & consistency.

(iii) Transaction :-

- smallest building block of a BC system that serves a purpose of BC.

(iv) chain :-

- sequence of blocks linked together using cryptographic hashes, forming an immutable chain.

(v) Blocks :-

- datastructures that store transactions.
- each block contain :

  Header – includes metadata

  Body – contains list of transactions.

(vi) Consensus mechanism :-

- set of rules and arrangements to carry out BC operations.
- common types
  - → Proof of Work (PoW) :
    Requires computational effort.
  - → Proof of stake (PoS) :
    Relies on stake of participants.
  - → Delegated Proof of stake (DPoS),
    Byzantine Fault Tolerance (BFT).

(vii) cryptography :-

- provides security and privacy
- utilize public & private keys to ensure data integrity & secure transactions.

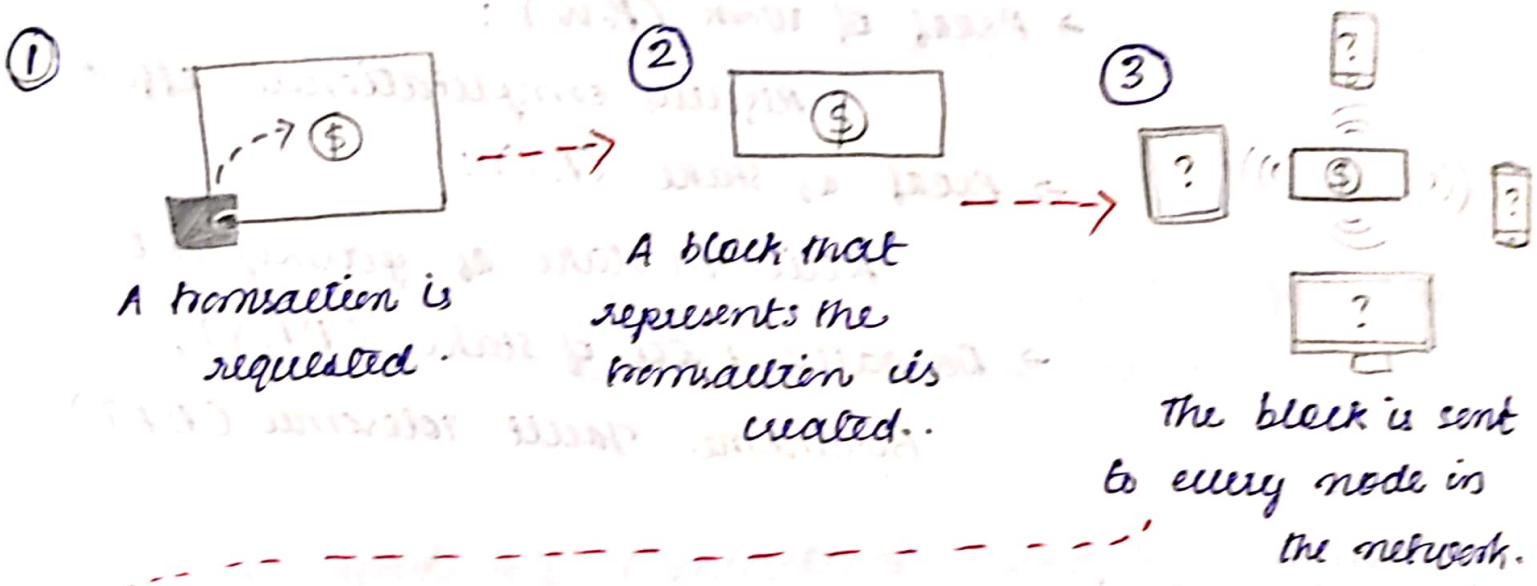
(viii) Miners :-

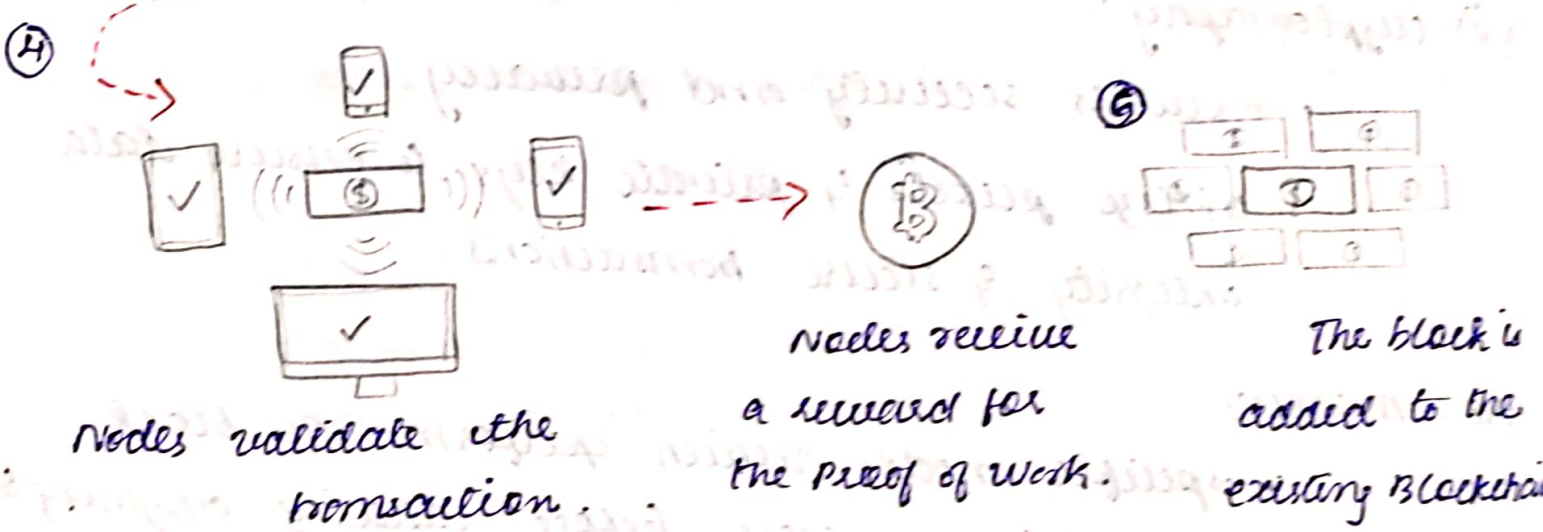
- specific nodes which perform the block verification process before adding anything to BC structure.

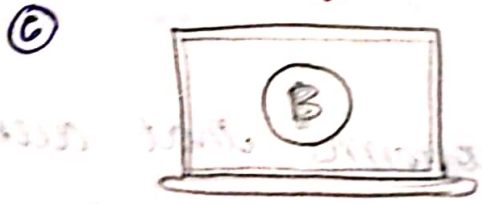
(ix) Smart contracts :-

- self executing programs that automatically carry out specific actions when certain conditions are met.

o) BC architecture diagram, that shows how this actually works in the form of a digital wallet.







o) Key characteristics of BC Architecture.

(i) Cryptography :

BC transactions are validated and trustworthy.

(ii) Immutability :

Any records made in a BC cannot be changed or deleted.

(iii) Provenance :

It is possible to track the origin of every transaction inside BC ledger.

(iv) Decentralisation :

each member of BC structure has access to the whole distributed database.

(v) Anonymity :

each BC network participant has a generated address, not user identity.

(vi) Transparency :

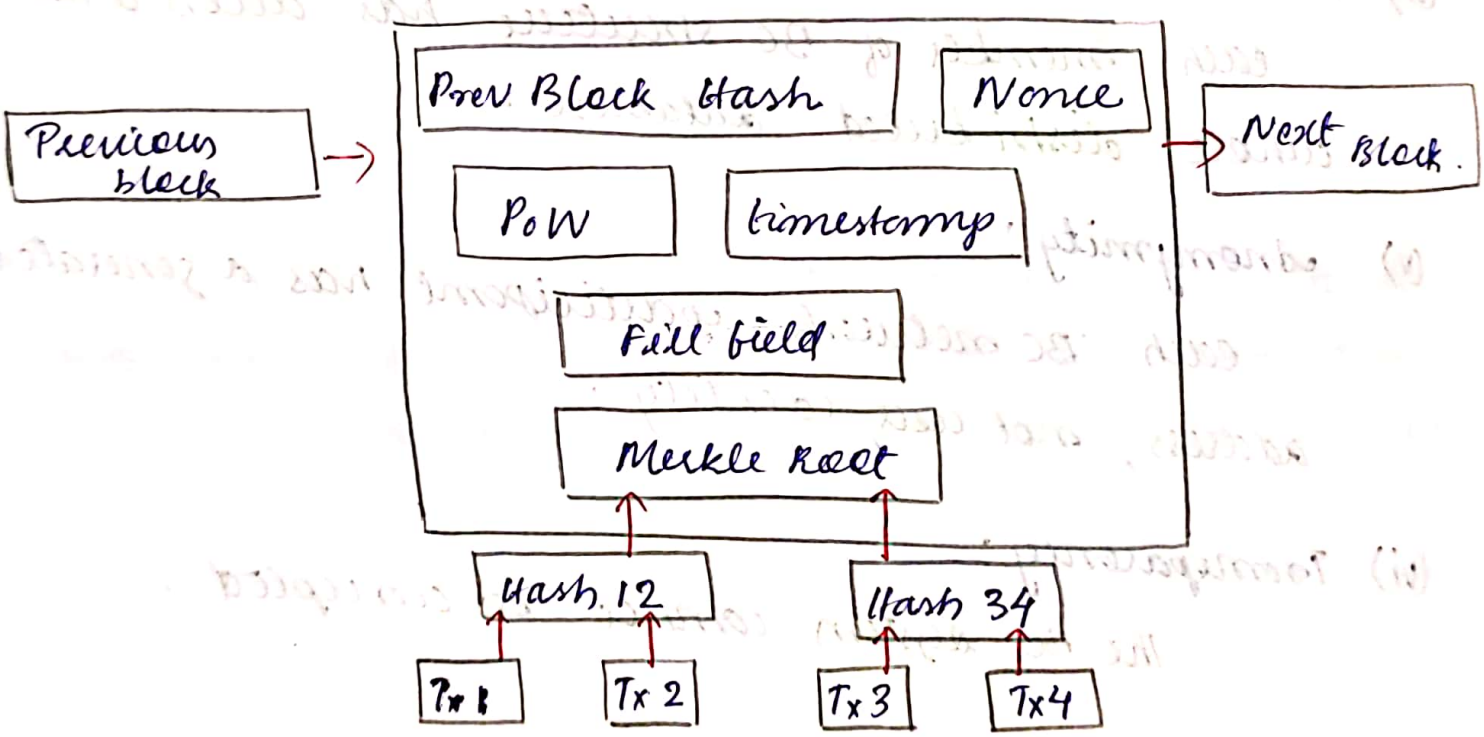
The BC system cannot be corrupted.

5.2 : Blockchain Data Structures

— are the fundamental building blocks that enable the secure, decentralized, nature of BC technology.

→ DS on different blockchain platforms are slightly different, but basically same.

→ Blockchain guarantees irreversible modifications of data, based on 2 hash structures, Merkle tree and block list.



1.

→ Merkle tree :

→ DS used in blockchain to organise & secure transaction data efficiently.

→ It helps verify and transfer large amounts

of data quickly in BC network.

→ every transaction on the BC is turned into a unique code called a hash.

→ instead of storing these hashes in straightline, they are arranged in a tree-like structure.

→ here's how it works:

- each transaction hash is placed at the bottom level of the tree, called leaf level.
- these hashes are paired & combined to create new hashes at the next level up.
- this process continues until only one hash is left at the top of the tree, called the merkle root.

→ for example, if there are 7 transactions in a block:

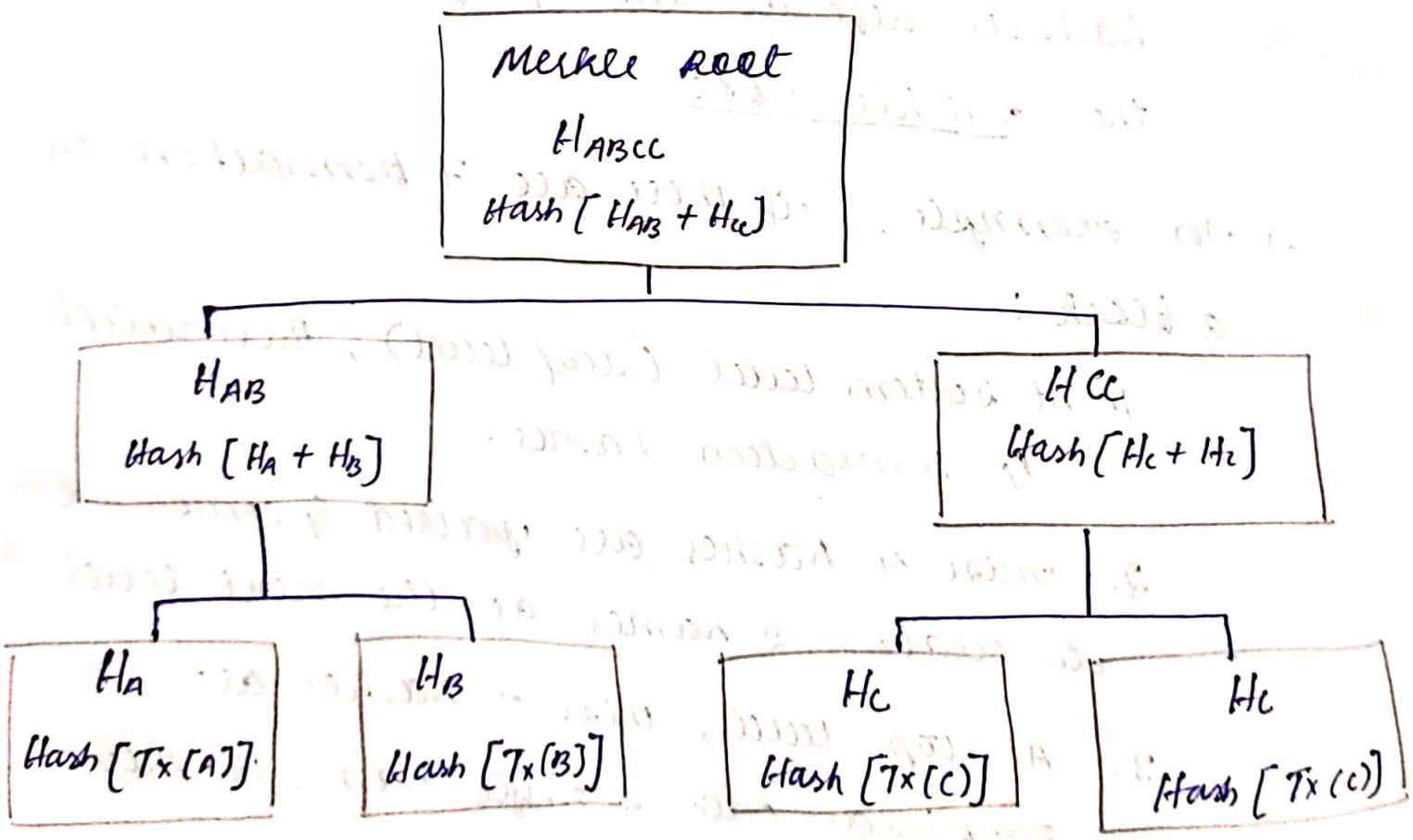
1. At bottom level (leaf level), there will be 4 transaction hashes.
2. These 4 hashes are paired & combined to create 2 hashes at the next level.
3. At top level, these 2 hashes are combined into a single hash, the merkle root.

→ This structure looks like an upside-down tree, where each node connects to 2 nodes below it.

→ The merkle root contains a summary of all the transaction data in the block.

→ It allows anyone to quickly verify if a transaction is a part of the block by checking its hash against merkle root, without needing to look at every transaction.

→ This makes BC secure and efficient for large amounts of data.



8. Blocklist :

[On a BC, each block contains imp. info in its block header, including the hash of prev. block, a special num called Nonce, and the Merkle root (which summarizes all transactions in block). To create hash value of a block, a special operation called SHA256 is performed.].

→ The block that comes after another is connected to it using PrevBlockHash, which stores hash of the previous block.

→ This creates a blocklist, where each block is linked to one before it.

→ The PrevBlockHash ensures that the blocks are ordered & cannot be changed without breaking the entire chain.

→ Bcoz each block contains the hash of the previous one, if someone tries to tampee with a block (changing its data), it will change the hash of that block, which in turn will change the names of all following blocks.

→ When blocks are downloaded from an untrusted source, you can use this hashing method to check if any blocks have been altered, ensuring the integrity & security of B.C.

3. Proof of Work:

→ is a system used to secure digital transactions without needing a trusted third party.

→ It works by having miners to solve a difficult mathematical puzzle to verify transactions and add new blocks to BC. This process is called mining.

→ And the first miner to solve the puzzle gets rewarded, often with new cryptocurrency like Bitcoin.

5.3: Blockchain Data types.

In BC systems, there are 2 primary types of data models used to track transactions and manage state of accounts.

1. Transaction-Based Data Model:

→ Here, a blockchain stores info primarily about transactions, rather than state of accounts.

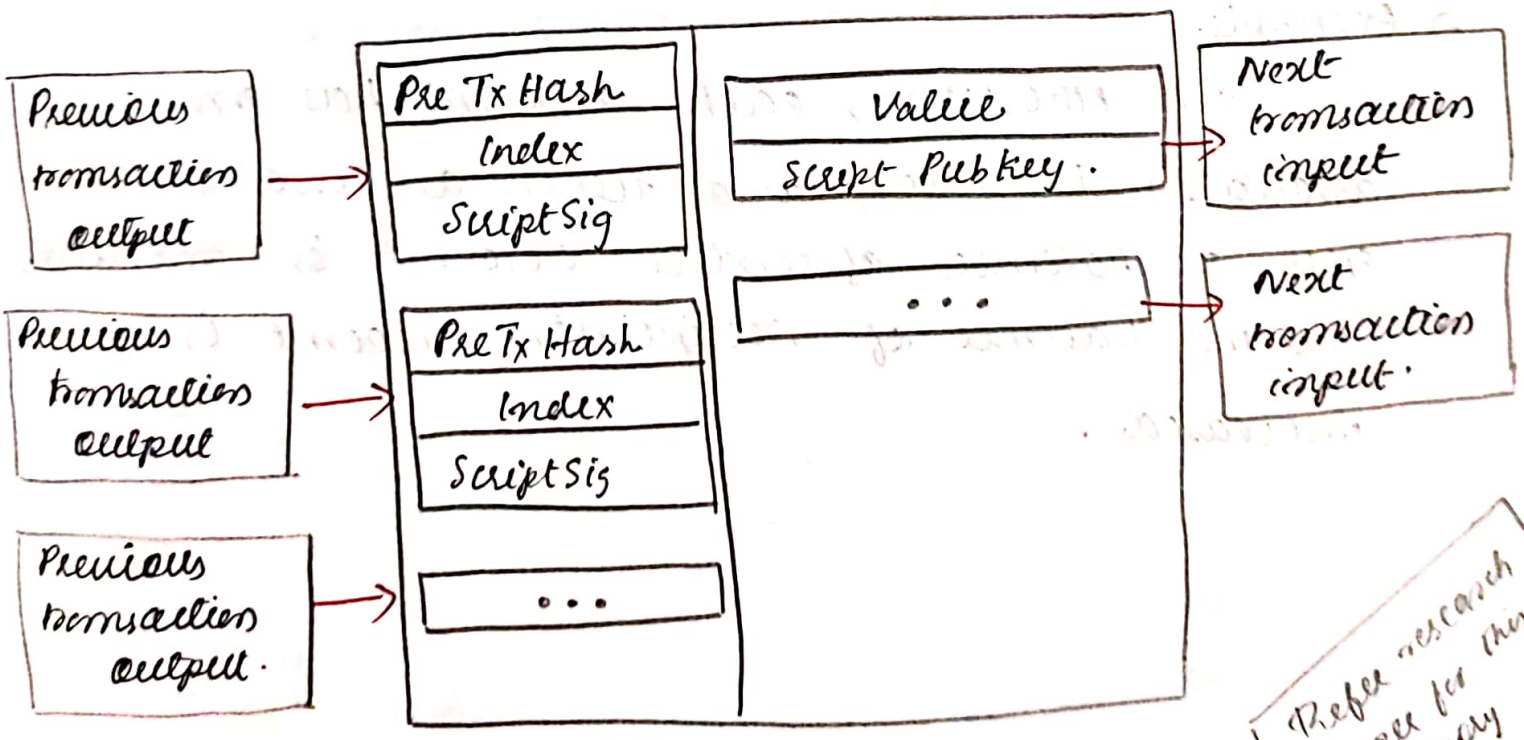
→ each transaction is recorded in sequence, & the BC maintains a ledger of all transactions that have occurred.

→ Popular BC like Bitcoin & Ethereum use this model.

→ How it works:

- In this model, when a transaction occurs, it is recorded as a new entry in the B.C.
- The BC doesnot store current balance of each account directly.
- Instead, it maintains a history of all transactions and derives the balance of an account by looking at all transactions related to that account.

→ Example: In Bitcoin, you can think of a user's balance as the total sum of inputs minus total sum of outputs.



2. Account Based Data Model

→ BC directly tracks the state of each account, including its current balance & other relevant data.

→ each account has a balance associated with it, and the system directly updates the balance whenever a transaction occurs.

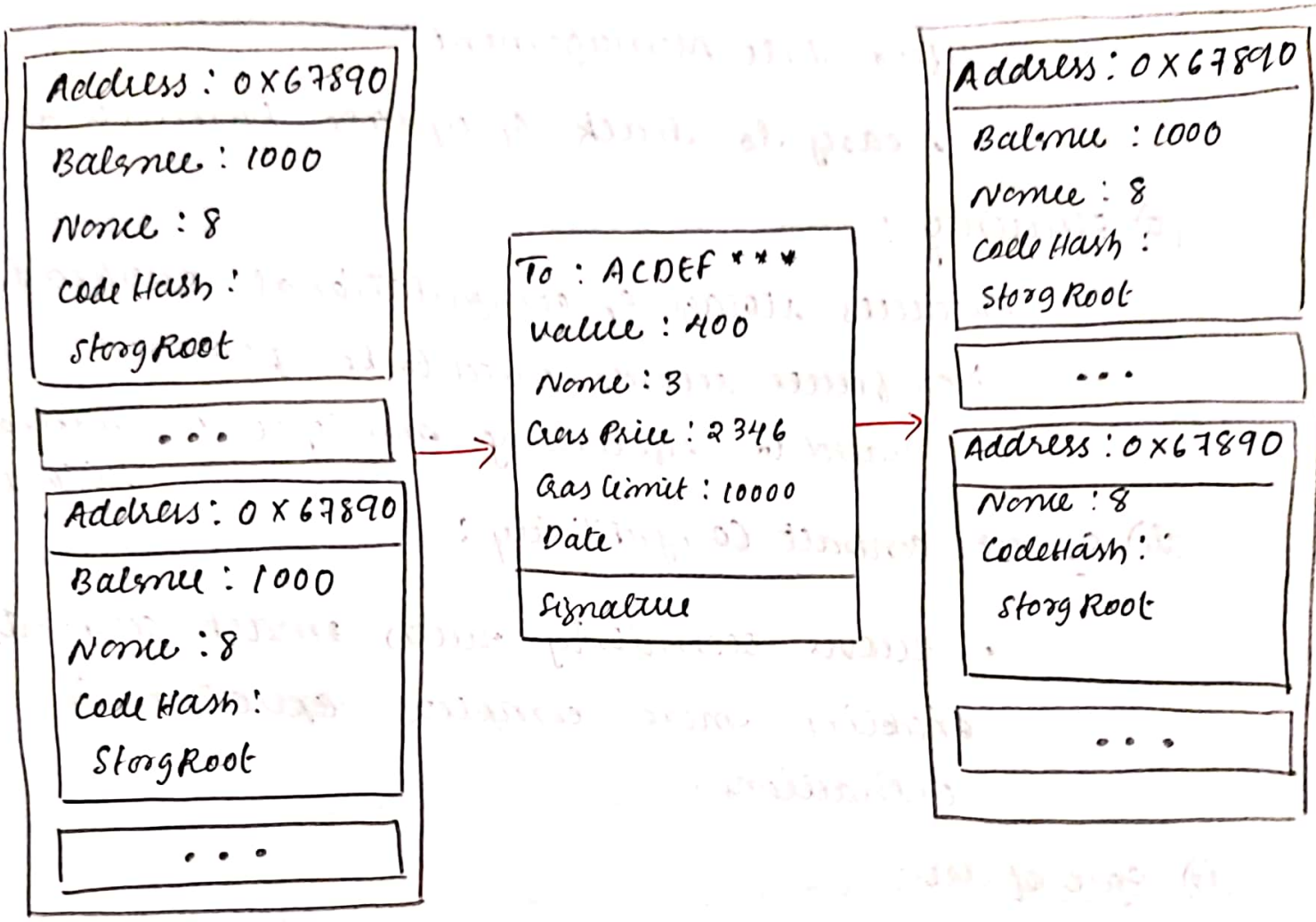
→ Ethereum uses this model, where each address has a balance that changes whenever a transaction is made.

→ How it works:

- Instead of having to look at the entire history of transactions to determine an account's balance, the system can simply check the current state of the account directly.

→ Example:

On Ethereum, each account has an associated balance, and when a user sends Ether, the balance of sender's account is reduced, and the balance of recipient's account is increased.



Benefits :

1. (i) High Transparency & Auditability
   - every transaction can be traced back to its origin, ensuring transparency & auditability of funds.

   (ii) Increased Privacy :
   - Addresses are not tied to an identity, new address can be created for each transaction, enhancing user privacy.

   (iii) Double-spending Prevention.
   - ensures each transaction o/p can only be spend once by verifying chain of ownership.

2. (i) Simplified State Management:

- easy to track & update balances directly

(ii) Efficiency:

- Reduces storage & computational overhead bcoz fewer records need to be processed compared to referencing multiple transaction histories.

(iii) Smart contract Compatibility:

- works seamlessly with smart contracts enabling more complex operations & interactions.

(iv) Ease of use:

- easier to develop applications & manage accounts since transactions involve straightforward balance updates.

5.4: Contract Data

→ In 1995, the first concept of Smart Contract was proposed.

→ A smart contract is a program on a BC that automatically does what its told when certain conditions are met.

→ Here's how it works:

- It's an automatic agreement:
  If 'x' happens, then 'y' will

automatically happen.

- No middlemen needed:
  It works on its own, so no third party like a bank or lawyer is required.
- Secure & transparent:
  Everything is recorded on the BC, so no one can change or cheat it.

→ Think of it like a vending machine. You put in money, press a button, and it gives you what you paid for automatically.

4) How do smart contracts work?

→ works by following simple if / when ... then... rules written into code on a B.C.

→ it runs automatically when certain conditions are met.

1. Conditions are set:
   The smart contract has clear rules like "If Buyer gets the product, then release payment".
2. Automatic execution:
   When the conditions are met, the contract automatically does what its supposed to.
3. Secure & tamper-proof:
   Everything happens on BC, so no one can change or cheat the system.

i) Benefits of smart contracts:

1. Speed & Accuracy:

- everything is automated & digital, so no more time wasted on paperwork or fixing manual mistakes.
- code is clear & precise, unlike complicated legal language.

2. Trust:

- smart contracts follow the rules automatically and everyone involved can see the encrypted transaction records.
- No one can secretly change the info for their own gain.

3. Security:

- data is encrypted & stored on BC, making it extremely hard to hack.
- To change one record, someone would have to alter the entire chain, which is nearly impossible.

4. Savings:

- No middlemen are needed, so you save on their fees.
- Participants can trust the system without paying for extra verifications.

- All tasks are handled by code, reducing costs.

.) How smart contracts work (new note).

1. Contract Generation :

- Negotiation : All parties discuss & agree on their roles & responsibilities.
- Specification : The agreed terms are written into a program.
- Verification : The program is tested to ensure it works as expected.
- Code creation : The final contract is turned into a standard code format.

2. Contract Release :

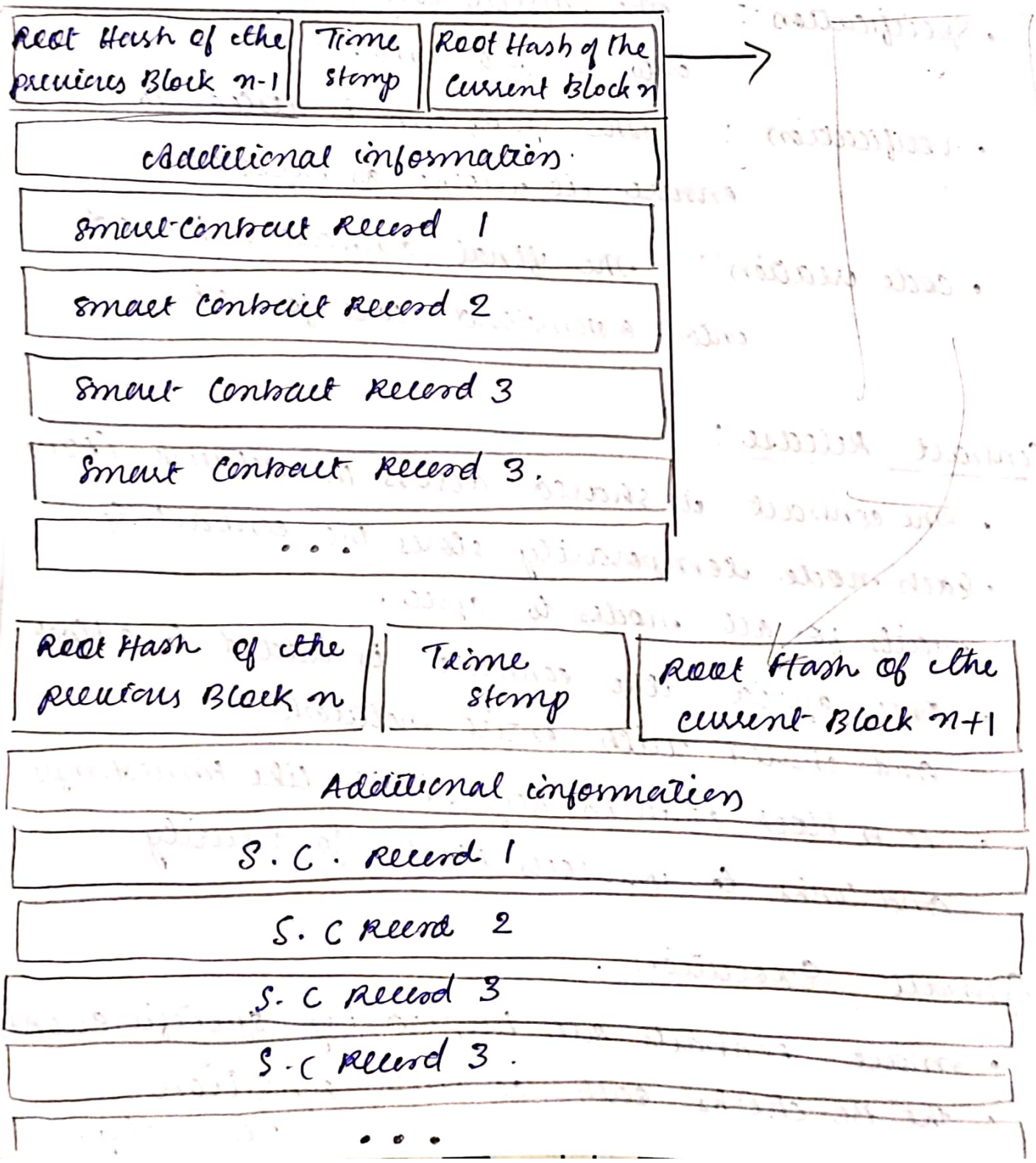
- The contract is shared across the network (P2P)
- Each node temporarily stores the contract & waits for all nodes to agree.
- Once agreed, the contract is added to a block, and shared with entire network.
- each block contains key details like timestamps and links to previous blocks for security.

3. Contract Execution :

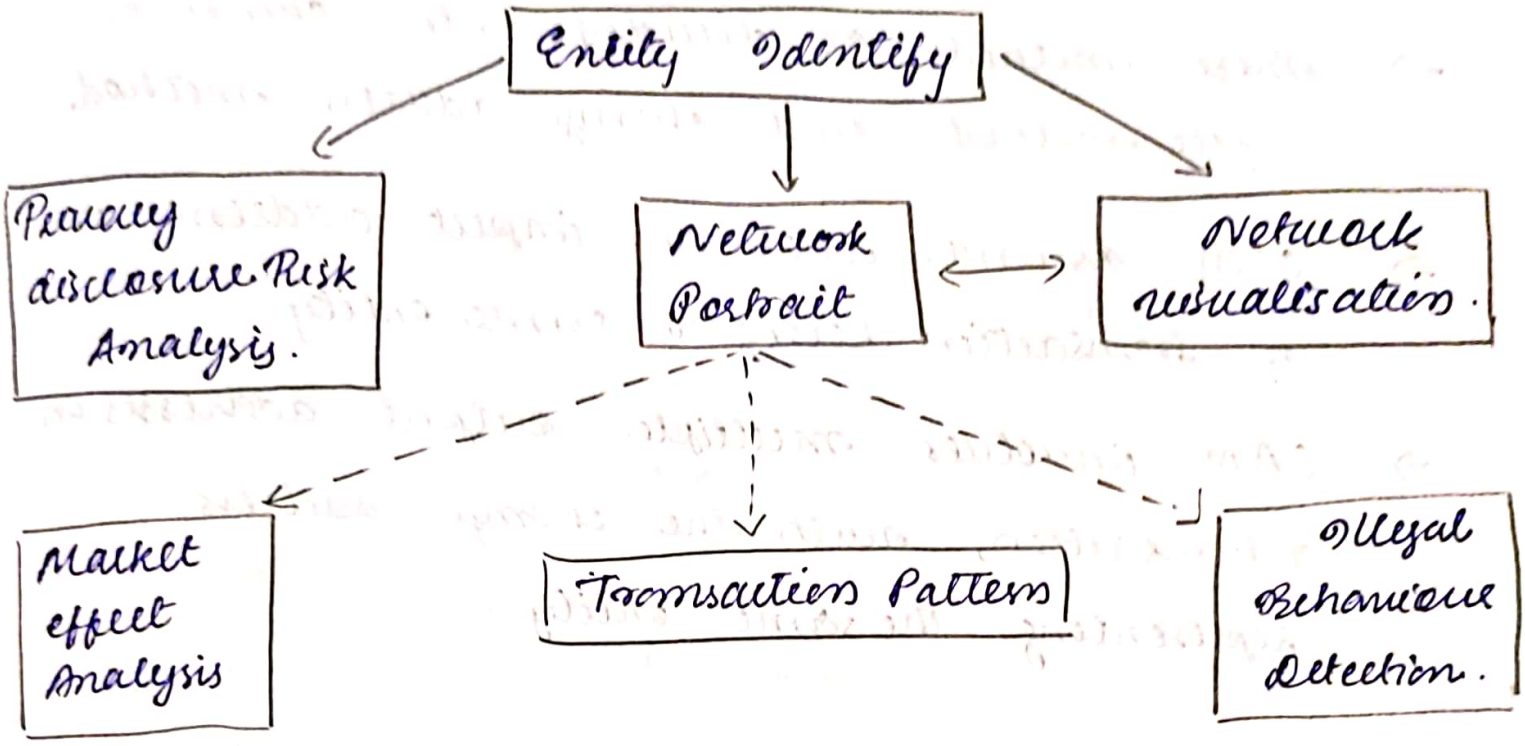
- Smart contracts are triggered by specific events.
- The BC checks each contracts conditions regularly.

- if the conditions are met, contract is pushed to a queue for verification.
- all nodes verify & agree on the contract before it is executed.

c) BC diagram of smart contract



5.5 Problems to be solved in B.C Data Analysis.



1. Entity Identification :

→ On BC transactions, users are anonymous.

→ Transactions involve multiple users, and a single user may participate in multiple transactions at some time.

→ Question is that : can you identify users from transaction records / determine which addresses belong to the same user?

→ since, it is not possible to confirm the identity of a user, it is considered that an entity is identified instead.

→ An entity may be a user / an organisation.

→ Heuristic methods are used to identify potential entities.

→ These methods are divided into common input method and change address method.

→ CIM assumes that all input addresses in a transaction belong to same entity.

→ CAM involves multiple output addresses in a transaction, with the change address representing the same entity.

2. Privacy Protection: (Block).

→ B.C Privacy protection has 2 main parts: Identity privacy & transaction privacy.

→ IPP ensure a users identity, physical address, and IP address are not linked to their public keys or addresses on BC.

→ TPP makes sure that transaction details (such as amount, sender's public key, recipients address, and purchase content) are kept private.

→ unauthorized users can't access transaction details through technical means.

3. Network Portrait :

→ Researches want to analyze how many users are involved in BC transactions.

→ They aim to understand characteristics of these users.

→ They want to know if BC's payment network behaves like a typical complex network.

→ They are interested in how BC is distributed among users.

→ They also want to know if BC follows general economic laws.

4. Network visualization :

→ BC technology is growing quickly, so is the amount of transaction data stored in it.

→ studying tools that help display and understand large & growing trading networks is an imp area of research.

→ one eg: is BitconeView, a system that helps track BC transactions real-time.

→ The system uses a concept called 'purity' to identify mixed currency transactions.

5. Market Effect Analysis :

→ Prices of cryptocurrencies like BC are very unstable.

→ This has led economists to debate whether BC can be considered a real currency.

→ Researchers want to understand what causes this extreme price change.

→ some believed factors include miner actions, how cryptocurrency is set up, how many people use it, govt rules, potential users, market mood & competition from other cryptocurrencies.

6. Illegal behaviour detection :-

→ BC is an anonymous & decentralised payment system, unlike traditional bank systems

→ Its anonymity leads to illegal activities like money laundering, fraud & illegal sales.

→ Identifying illegal behaviours through transaction patterns & blockchain data analysis is imp.

→ This helps promote healthy development of blockchain technology

7. Transaction pattern recognition :

→ Bitcoin is an anonymous & decentralised payment system, unlike traditional bank systems.

→ An interesting question is how human payment behaviour changes in an anonymous system.

→ It is valuable to know specific patterns can be found in Blockchain transactions to detect illegal activities.