MODULE 3

Hashing in Blockchain

Blockchain technology is an intricate web of several technological innovations working together. Among the most important pieces of the blockchain puzzle is hashing.

Hashing is a cryptographic function that converts a string of characters of any length into a unique output, or hash, of fixed length. This means that no matter what combination of symbols are used as the input, they will always produce a one-of-a-kind string of digits and characters.

A Bitcoin hash looks like this:

00000000000000000025e2ba026a8ad462b9a693d80fd0887def167f5f888a11

(hash of block 540807)

i. Hashing Essentials

- Hashing is a method for cryptographically encoding data.
- It produces a fixed-length output from any input.
- The same input always produces the same hash.
- The input cannot be reconstructed from the hash.
- Modern hash functions make it virtually impossible to produce the same output from two different inputs.

ii. Hashing in Cryptocurrencies

Hashing is an integral part of all blockchain-based transactions, including the trading of cryptocurrency. Hash functions are necessary in everything from mining blocks to signing transactions to generating private keys.

Bitcoin	SHA-256
Bitcoin Cash	SHA-256
Ethereum	keccak256
Litecoin	scrypt
Ripple	SHA-512

A hash function is a mathematical algorithm used to calculate the hash. Different cryptocurrencies use different hash functions but all of them follow the same basic principles of hashing.

iii. Main Properties of Hashing

- **Hashing produces outputs of fixed length**

Hashing will always produce a unique, fixed-length output from any input. Let us take a look at what that means with a couple of examples.

Input: hello

Output: 2CF24DBA5FB0A30E26E83B2AC5B9E29E1B161E5C1FA7425E73043362938B9824

Input: It’s a good day to HODL

Output: 6B89D5D4AD6A3364410DD9BAB95FD250EF4A663D9D3C47CBD7388535A5912E03

Input: The entire novel of Oliver Twist

Output: 4F144CC612CA27E2DD6DFD6663F68BABC3B758D602B5102BF14E717E823EB741

Here, the SHA-256 hash function is used to generate the hashes of two different inputs. In all three cases, the hash is completely unique, but its length remains the same. SHA-256 generates hashes that are 256 bits long, usually represented as 64 symbols comprised of numbers 0–9 and letters A–F. No matter how short or how long the input is – be it a single word (hello) or even a whole novel (Oliver Twist by Charles Dickens) – the hash is fixed at 64 characters.

- **Hashing is deterministic**

The same input will always produce the same output. If you use SHA-256 to generate a hash from "fun", you will always get the output seen in the table below. Even changing one letter, however, will produce a completely different hash.

Input: fun

Output: 00C4285274FCC5D6FBA2EE58DAF0D8C2B9B825B68D35D65D0E90A9BB333A51B5

Input: sun

Output: 27756F050E14A1CB1C1EE867F0EACE9EA4D9FCB81B8BEE089469F1EBD5FD7B17

- **Hashing is a one-way function**

It is infeasible to determine what the input was from any given output. That is to say, it is virtually impossible to reverse the hash function with contemporary technology. The only way to determine what the input was is trying out random strings until you find the right one. This method is known as brute force.

Using brute force to reverse the hash back to the original string is easier said than done. No computer in existence is powerful enough to find the solution in any reasonable amount of time, nor are we ever likely to build one that will. Even IBM Summit, currently the fastest computer in the world, capable of making several trillion calculations per second, would need many years and an astounding amount of electricity to find the answer for a single hash.

- **Hashing is resistant to collisions**

A collision occurs when a hashing mechanism produces the same output for two different inputs. This is possible in theory for hashing, as the number of unique hashes is limited but the number of inputs is not. However, the probability of collisions is extremely small. Hashing is thus said to be resistant, but not immune, to collisions.

SHA-256, the algorithm used by Bitcoin, outputs hashes that are 256 bits long (a 256 digit-long string of 1s and 0s). This means there are a total of 2^{256} unique hashes that it can produce. As soon as the number of inputs is larger than the number of all possible outputs, let us say $2^{256}+1$, at least two of the inputs will have the same output – that's a collision.

So does that mean hashing is exploitable? No, not at all. 2^{256} is an enormous number. The sheer size of this number means the likelihood of a collision occurring is utterly miniscule.

Concept of Double Spending

Double-spending is the risk that a cryptocurrency can be used twice or more. Transaction information within a blockchain can be altered if specific conditions are met. The conditions allow modified blocks to enter the blockchain; if this happens, the person that initiated the alteration can reclaim spent coins.

To understand double-spending, it helps to review how the blockchain works first. When a block is created, it receives a hash—or encrypted number—that includes a timestamp, information from the previous block, and transaction data. This information is encrypted using a security protocol like the SHA-256 algorithm used by Bitcoin. Once that block's information is verified by miners (in proof-of-work consensus), it is closed, and a new one is created with the timestamp, transaction information, and previous block's hash. A Bitcoin is awarded to the miner whose machine verified the hash.

For someone to double spend, a secret block has to be mined that outpaces the creation of the real blockchain. They would then need to introduce that chain to the network before it caught up—if this happened, then the network would recognize it as the latest set of blocks and add it to the chain. The person that did this could then give themselves back any cryptocurrency they had spent and use it again.

- **Preventing Double Spending**

Double spending remains a risk; however, it is minimized by the blockchain. The likelihood of a secret block being inserted into the blockchain is very slim because it has to be accepted and verified by the network of miners. The only chance a miner with illicit intentions has of inserting an altered block is to attempt to get another user to accept a transaction using their secret block and cryptocurrency. Even then, the likelihood that the modified block will be accepted is very slim. The blockchain and consensus mechanism move so quickly that the modified block would be outdated before it was accepted. Even if it was accepted, the network would still have passed up the information in the block and would reject it. Cryptocurrency transactions take some time to verify because the process involves randomly selecting numbers to solve the complex hash—this also takes up a great deal of computational power. It is, therefore, exceedingly difficult to duplicate or falsify the blockchain because of the immense amount of computing power needed to stay ahead of all of the other miners on the network.

- **Double Spending Attacks**

The most significant risk for blockchains comes in the form of a 51% attack, which can occur if a miner controls more than 50% of the computing power that validates the transactions, creates blocks, and awards cryptocurrency. If this user—or users—controls a majority of the hashing in the blockchain, they will be able to dictate transaction consensus and control the award of currency. In more popular cryptocurrencies such as Bitcoin, this is very unlikely due to the number of miners and hashing difficulty it has reached; however, new or forked cryptocurrencies with smaller networks are susceptible to this attack. Most commonly, the unconfirmed transaction attack is used to fool cryptocurrency users. If you see one of these transactions, you shouldn't accept it because it can cause an attempted double-spend attack.

Mining

Blockchain mining is a peer-to-peer computer process and is used to secure and verify bitcoin transactions. Mining involves blockchain miners who add bitcoin transaction data to Bitcoin's global public ledger of past transactions. In the ledgers, blocks are secured by blockchain miners and are connected to each other forming a chain.

When we talk in-depth, as opposed to traditional financial services systems, Bitcoins have no central clearing house. Bitcoin transactions are generally verified in decentralized clearing systems wherein people contribute computing resources to verify the same. This process of verifying transactions is called mining. It is probably referred to as mining as it is analogous to mining of commodities like gold—mining gold requires a lot of effort and resources, but then there is a limited supply of gold; hence, the amount of gold that is mined every year remains roughly the same. In the same manner, a lot of computing power is consumed in the process of mining bitcoins. The number of bitcoins that are generated from mining dwindles over time. In the words of Satoshi Nakamoto, there is only a limited supply of bitcoins. Only 21 million bitcoins will ever be created.

At its core, the term 'Blockchain mining' is used to describe the process of adding transaction records to the bitcoin blockchain. This process of adding blocks to the blockchain is how transactions are processed and how money moves around securely on Bitcoins. This process of blockchain mining is performed by a community of people around the world called 'Blockchain miners.'

Anyone can apply to become a blockchain miner. These Blockchain miners install and run a special blockchain mining software that enables their computers to communicate securely with one another. Once a computer installs the software, joins the network, and begins mining bitcoins, it becomes what is called a 'node.' Together, all these nodes communicate with one another and process transactions to add new blocks to the blockchain which is commonly known as the bitcoin network. This bitcoin network runs throughout the day. It processes equivalent to millions of dollars in bitcoin transactions and has never been hacked or experienced downtime since its launch in 2009.

i. Types of Mining

The process of mining can get really complex and a regular desktop or PC cannot cut it. Hence, it requires a unique set of hardware and software that works well for the user. It helps to have a custom set specific to mining certain blocks.

The mining process undertaking can be divided into three categories:

a. Individual Mining

When mining is done by an individual, user registration as a miner is necessary. As soon as a transaction takes place, a mathematical problem is given to all the single users in the blockchain network to solve. The first one to solve it gets rewarded. Once the solution is found, all the other miners in the blockchain network will validate the decrypted value and then add it to the blockchain thus verifying the transaction.

b. Pool Mining

In pool mining, a group of users works together to approve the transaction. Sometimes, the complexity of the data encrypted in the blocks makes it difficult for a user to decrypt the encoded data alone. So, a group of miners works as a team to solve it. After the validation of the result, the reward is then split between all users.

c. Cloud Mining

Cloud mining eliminates the need for computer hardware and software. It's a hassle-free method to extract blocks. With cloud mining, handling all the machinery, order timings, or selling profits is no longer a constant worry. While it is hassle-free, it has its own set of disadvantages. The operational functionality is limited with the limitations on bitcoin hashing. The operational expenses increase as the reward profits are low. Software upgrades are restricted and so is the verification process.

ii. Mining Bitcoins

a. Mining Bitcoins in Cloud

Obtain a bitcoin wallet: Bitcoins are stored in digital wallets in an encrypted manner. This will keep your bitcoins safe.

Secure the wallet: Since there is no ownership of bitcoins, anyone who gains access to your blockchain wallet can use it without any restriction. So, enable two-factor authentication and store the wallet on a computer that does not have access to the Internet or store it on an external device.

Choose a cloud mining service provider: Cloud mining service providers allow users to rent processing or hashing power to mine bitcoins remotely. Popular cloud mining service providers are Genesis Mining and HashFlare.

Choose a cloud mining package: To choose a package, you will need to decide on how much you are willing to pay and keep your eyes open to the hashing power the package will offer. Cloud mining companies will mostly envisage the Return on Investment (ROI) based on the current market value of Bitcoins.

Pick a mining pool: This is the best shot you can get to earn bitcoins easily. There are many mining pools which charge a mere 2 percent of your total earnings. Over here, you will have to create workers which are basically subaccounts that can be used to track your contributions to the pool.

Put your earnings in your own secure wallet: Whenever you witness an ROI, simply withdraw your earnings and put them in your own secure wallet.

b. Mining Bitcoins on your own

Purchase custom mining hardware: You need to purchase an Application-specific Integrated Circuit (ASIC) miner to mine bitcoins. While purchasing an ASIC Blockchain miner, you should consider its efficacy in hashing power and take note of its pricing policies.

Purchase a power supply: Blockchain miners consume a lot of power. So, get a dependable power supply that is compatible with the ASIC miner that you purchase.

Obtain a bitcoin wallet: Bitcoins are stored in digital wallets in an encrypted manner. This will keep your bitcoins safe.

Secure the wallet: Since there is no ownership on bitcoins, anyone who gains access to your wallet can use it without any restriction. So, enable two-factor authentication and store the wallet on a computer that does not have access to the Internet or store it in an external device.

Pick a mining pool: This is the best shot you can get to earn bitcoins easily. There are many mining pools that charge a mere 2 percent of your total earnings. Over here, you will have to create workers which are basically subaccounts that can be used to track your contributions to the pool.

iii. Uses of Blockchain Mining

a. Validating Transactions

Bitcoin transactions take place in huge figures every day. Cryptocurrencies function without a central administrator and the insecurity can be substantial with the transactions that transpire. So, what is the authentication method with such cryptocurrencies? With each transaction, new blocks are added to the blockchain in the network and the validation lies in the mining results from the blockchain miners.

b. Confirming Transactions

Miners work the blockchain mining process to confirm whether the transaction is authentic or not. All confirmed transactions are then included in the blockchain.

c. Securing Network

To secure the transaction network, bitcoin miners work together. With more users mining the blockchain, blockchain network security increases. Network security ensures that there are no fraudulent activities happening with cryptocurrencies.

Proof of Work (PoW)

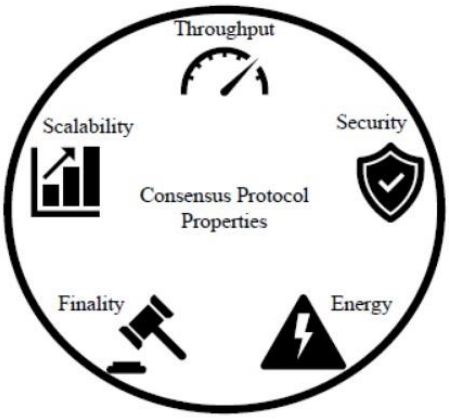
Proof-of-Work (PoW) in the blockchain is a consensus mechanism that allows miners to add a new block to the network based on calculations made to find the perfect hash. Network participants verify the transactions added by the new block.

For a decentralized network like Blockchain, keeping all network participants in sync is essential. However, it seems far-fetched for everyone to agree on one thing. Blockchain uses a consensus mechanism to create governance among all network participants.

i. Consensus Mechanism

Consensus means reaching a decision that all network participants agree on. For example, a group of friends agrees to play soccer without conflict. Reaching a decision to play football together is a state of consensus or mutual agreement here.

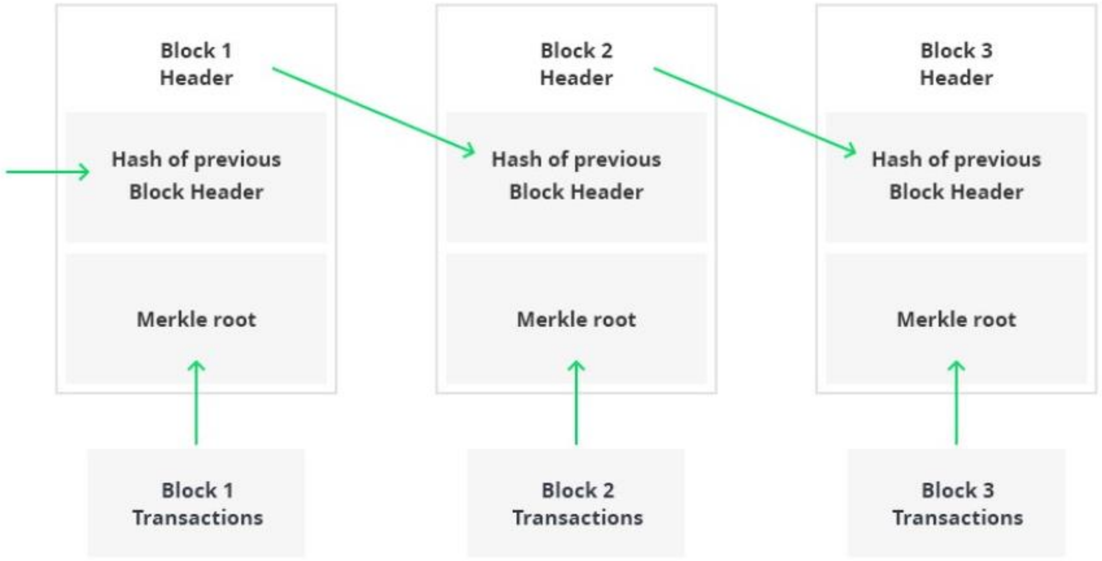
In the case of Blockchain, at least 51% of nodes or network participants agree on the upcoming change. If this happens, the network is updated with the new change. Otherwise, they reject the change by mutual agreement.



The Proof-of-Work (PoW) consensus mechanism is the oldest yet most popular. The idea emerged in 1993 when Moni Naor and Cynthia Dowrk published a paper exploring the potential of algorithms to prevent fraud.

PoW plays a significant role in the development of Blockchain technology. The goal is to create an authentication system that is hard to crack. The decentralized network works on the principle of distrust but cooperation. Blockchain (decentralized network) is a chain of linearly connected blocks containing information secured by cryptography. Here, each block contains the hash of its previous block to stay connected.

Additionally, each block contains several other information such as timestamp, block height, transaction records, Merkle Root Hash, block hash, previous block hash, difficulty level, and many more in the block header. The second part contains a set of financial transactions, the hash of which is eventually converted to a Merkle root. So Blockchain is defined as the chain of blocks of transactions.



Adding a new block to the chain is considered a new update to the current system. It, therefore, requires permission from network participants. In order to decide whether to add a new block or not, Proof-of-Work (PoW), a consensus mechanism, is used. Only verified transactions are added to the network.

In contrast, not all blocks are valid. Most proposed block networks are considered invalid. The Blockchain protocol defines the validity of a block. The blockchain network has an arbitrary "Difficulty" setting managed by the protocol that changes how hard it is to mine a block. Mining here means adding a new block.

Miners design new blocks in the chain. They are externals who want to add their block to the network. The work required to develop a valid block is where the value comes from. Miners receive rewards in proportion to their share of the computing power they spend mining a new block. The miner proves the work done by mining a valid block.

The difficulty level can vary in blockchains such as the Bitcoin network or Ethereum to ensure that blocks are created at regular intervals.

ii. How does PoW algorithm work?

The Proof-of-Work (PoW) consensus algorithm works by requiring each miner to overcome a difficulty level to prove the validity of a block. A block is marked as "valid" only if the hash value of the entire block is lower than the difficulty hash.

Block Hash < Hash Difficulty

The block contains important transaction information that cannot be changed. So miners change the nonce to get the hash below the difficulty threshold. A nonce is a part of a block that can be changed to limit the difficulty level.

Let's take an example to understand the working of PoW

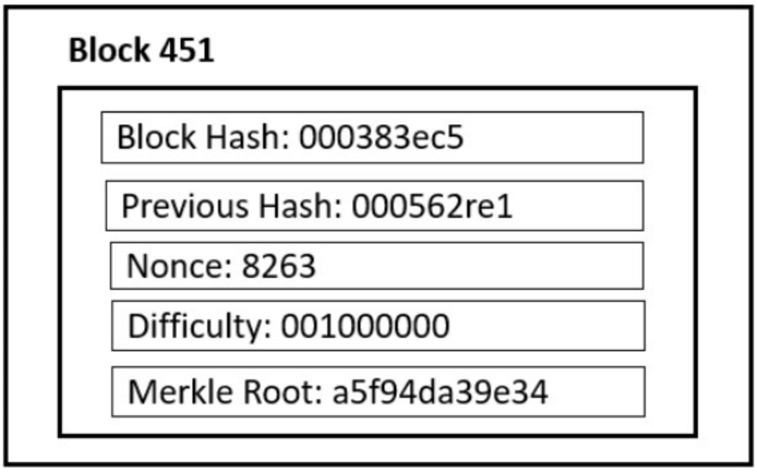
Harry is a Bitcoin miner who wishes to add his block of Bitcoin (digital currency) transactions to the network. However, for his block to be valid, he must change the block until the hash of his block falls below the difficulty threshold.

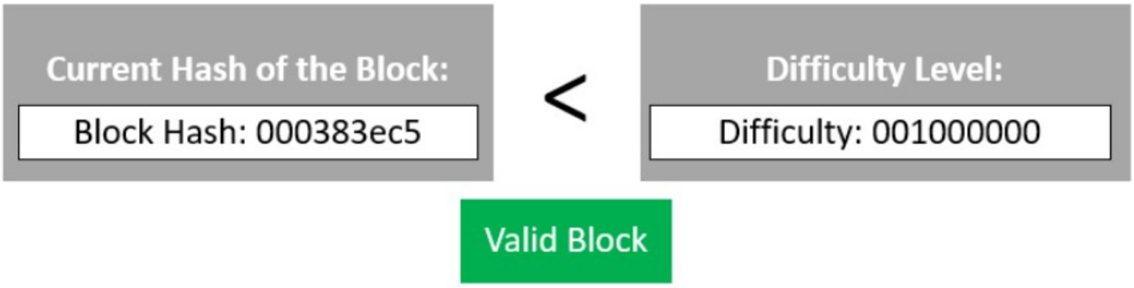
Let's say:

Harry block Hash: 817de9e0c

Hash difficulty: 001000000

Nonce: 8263





Harry will change the nonce until he gets the first 3 digits as zeros.

After continuously changing the nonce for hours, he finally got the hash.

Harry block Hash: 000383ec5

Hash difficulty: 001000000

Nonce: 6778

The difficulty threshold has now been reached. **Block Hash < Hash Difficulty.**

Therefore, Harry's block will be marked as valid and added to the blockchain. For mining a block in the Bitcoin blockchain, Harry gets a few Bitcoins as a block reward for spending the computing power to find a valid hash.

This process is completely based on chance. So the miner's job is to change the nonce value until the total hash of the block is lower than the difficulty hash.

iii. Advantages of Proof-of-Work

Below are the advantages of the Proof-of-Work (PoW) mechanism:

- A hard-to-find solution. Still, easy verification.
- As an initial consensus mechanism, PoW does not need initial stakes of coins before mining. One can start with 0 coins and it will only be positive.
- Ease of implementation compared to other blockchain consensus mechanisms.
- It is fault tolerant. It means that the failure of one component will not shut down the entire blockchain network.
- Give miners the opportunity to earn by adding a block.
- PoW is the oldest, most trusted, and most popular consensus protocol.

iv. Limitations of Proof-of-Work

- A lot of energy is wasted because only one miner can finally add their block.
- It requires a lot of computing power and, therefore, massive consumption of resources and energy.
- 51% risk of network attack. A controlling person can get 51% to control the network.
- Spread environmental hazards with attachment machines.
- PoW is a time and energy wipe-out process.
- It required a lot of hardware costs.
- Risk of Denial of Service Attacks by Intruders.

v. Proof of Work vs. Proof of Stake

Proof of work and Proof of stake are two discrete consensus mechanisms for cryptocurrency, but there are important differences between them.

Both methods confirm incoming transactions and add them to the blockchain. With Proof of Stake, network participants are known to as "validators" other than miners. One important difference is that instead of solving math problems, validators lock up a set amount of cryptocurrency – their stake – in a smart contract on the blockchain.

In interchange for "staking" cryptocurrency, they get a chance to prove new transactions and earn a reward. However, if they incorrectly verify wrong or fraudulent data, they may lose some or all of their deposit as a penalty.

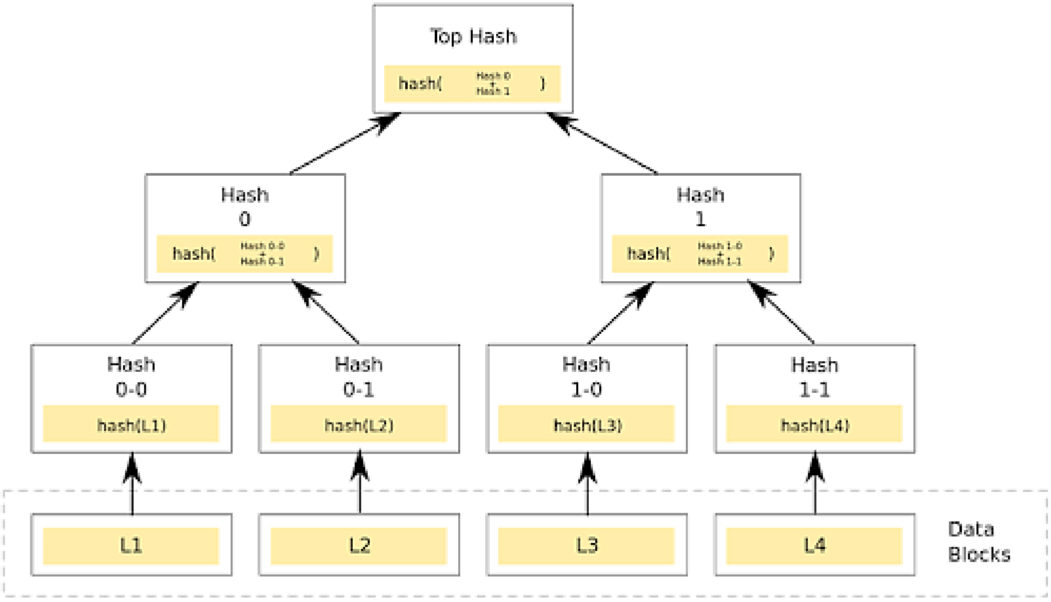
Proof of stake makes it easy to involve more people in blockchain systems as validators. There is no need to buy expensive computing systems and consume huge amounts of electricity to bet cryptocurrencies. All you need are coins.

Introduction to Merkel Tree

A hash tree, also known as a Merkle tree, is a tree in which each leaf node is labelled with the cryptographic hash of a data block, and each non-leaf node is labeled with the cryptographic hash of its child nodes' labels. The majority of hash tree implementations are binary (each node has two child nodes), but they can also have many more child nodes.

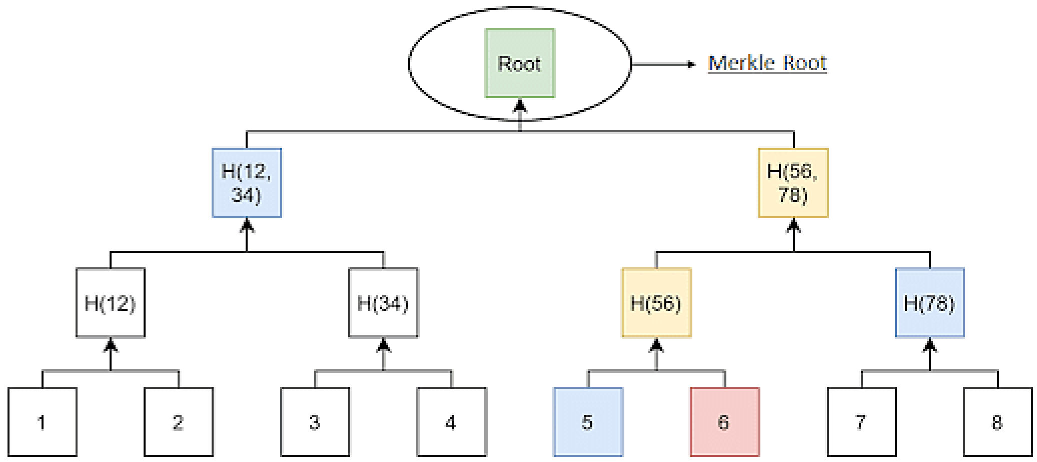
i. What is a Merkel Tree?

- Merkle trees, also known as Binary hash trees, are a prevalent sort of data structure in computer science.
- In bitcoin and other cryptocurrencies, they're used to encrypt blockchain data more efficiently and securely.
- It's a mathematical data structure made up of hashes of various data blocks that summarize all the transactions in a block.
- It also enables quick and secure content verification across big datasets and verifies the consistency and content of the data.



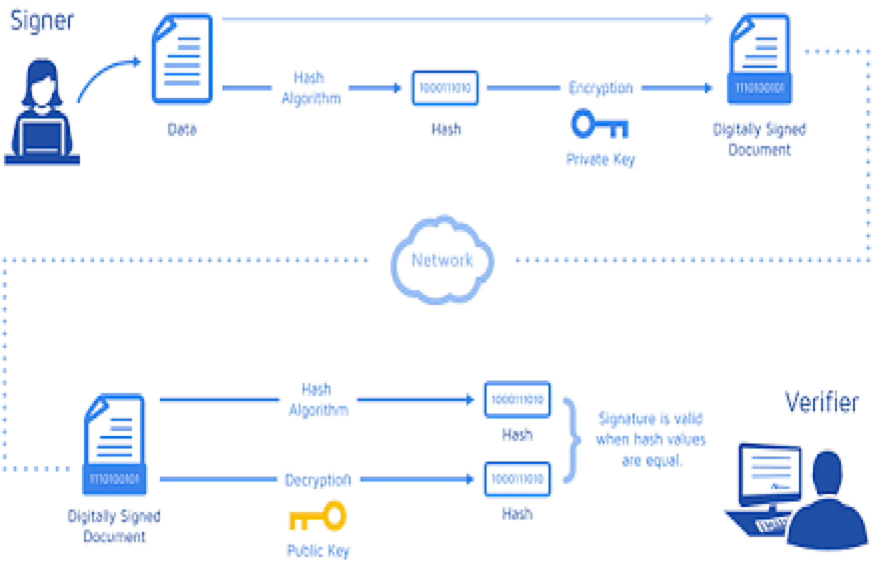
ii. What is a Merkel Root?

- A Merkle root is a simple mathematical method for confirming the facts on a Merkle tree.
- They're used in cryptocurrency to ensure that data blocks sent through a peer-to-peer network are whole, undamaged, and unaltered.
- They play a very crucial role in the computation required to keep cryptocurrencies like bitcoin and ether running.



iii. Working of Merkel Trees

A Merkle tree totals all transactions in a block and generates a digital fingerprint of the entire set of operations, allowing the user to verify whether it includes a transaction in the block.



Merkle trees are made by hashing pairs of nodes repeatedly until only one hash remains; this hash is known as the Merkle Root or the Root Hash. They're built from the bottom, using Transaction IDs, which are hashes of individual transactions. Each non-leaf node is a hash of its previous hash, and every leaf node is a hash of transactional data.

Example:

Consider the following scenario: A, B, C, and D are four transactions, all executed on the same block. Each transaction is then hashed, leaving you with:

- Hash A

- Hash B
- Hash C
- Hash D

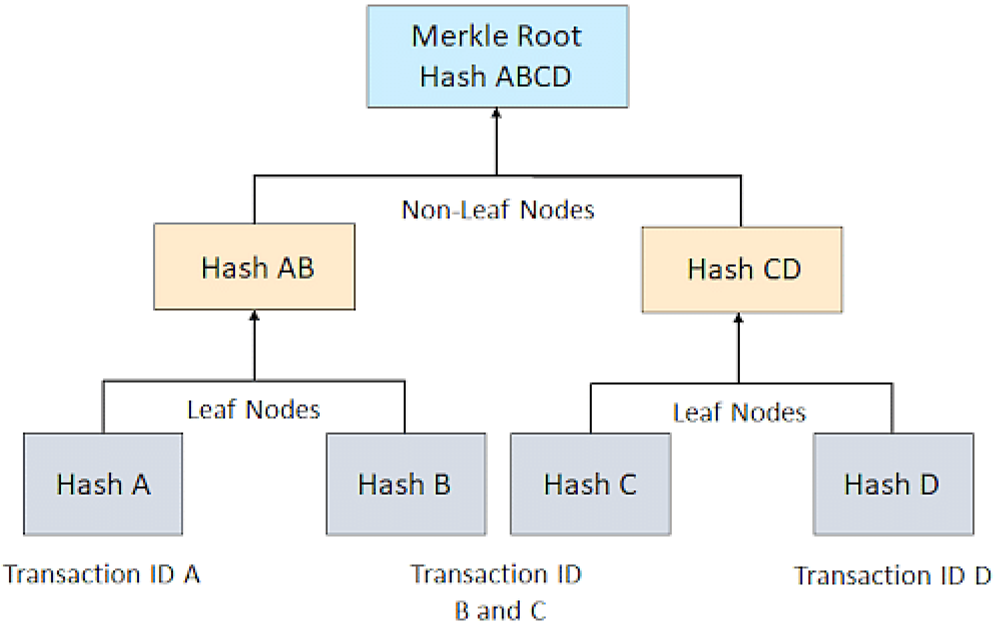
The hashes are paired together, resulting in:

- Hash AB

and

- Hash CD

And therefore, your Merkle Root is formed by combining these two hashes: Hash ABCD.



In reality, a Merkle Tree is much more complicated (especially when each transaction ID is 64 characters long). Still, this example helps you have a good overview of how the algorithms work and why they are so effective.

iv. Benefits of Merkle Tree in Blockchain

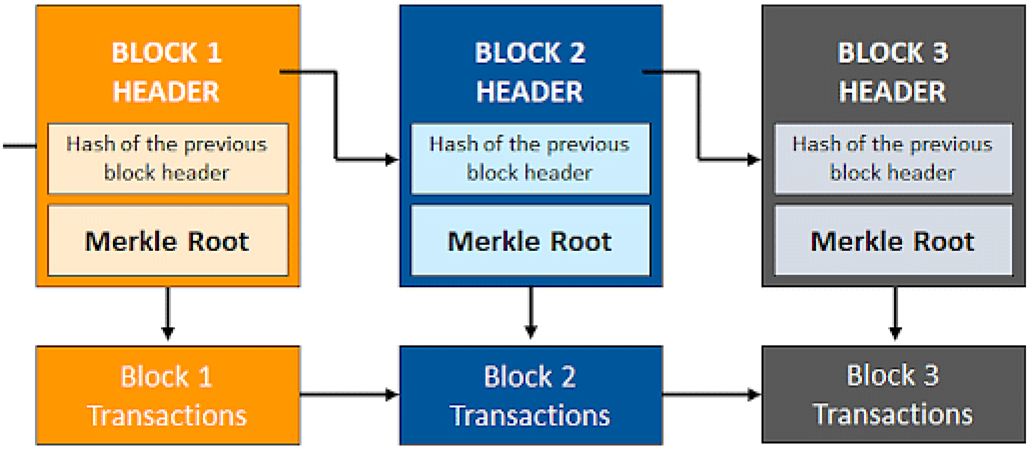
Merkle trees provide four significant advantages -

- ***Validate the data's integrity:*** It can be used to validate the data's integrity effectively.
- ***Takes little disk space:*** Compared to other data structures, the Merkle tree takes up very little disk space.
- ***Tiny information across networks:*** Merkle trees can be broken down into small pieces of data for verification.
- ***Efficient Verification:*** The data format is efficient, and verifying the data's integrity takes only a few moments.

v. Why is it essential to Blockchain?

Think of a blockchain without Merkle Trees to get a sense of how vital they are for blockchain technology. Let's take Bitcoin scenario because its use of Merkle Trees is essential for the cryptocurrency.

If Bitcoin didn't include Merkle Trees, per se, every node on the network would have to retain a complete copy of every single Bitcoin transaction ever made. One can imagine how much information that would be. Any authentication request on Bitcoin would require an enormous amount of data to be transferred over the network: therefore, you'll need to validate the data on your own. To confirm that there were no modifications, a computer used for validation would need a lot of computing power to compare ledgers.



Merkle Tree breaking the data into tiny parts of information

Merkle Trees are a solution to this issue. They hash records in accounting, thereby separating the proof of data from the data itself. Proving that giving tiny amounts of information across the network is all that is required for a transaction to be valid. Furthermore, it enables you to demonstrate that both ledger variations are identical in terms of nominal computer power and network bandwidth.

Data Privacy and Blockchain

Data Privacy is sometimes referred to as information privacy, which deals with the proper handling of sensitive data including personal data. Data privacy has regulated the manner in which personal data is collected, processed, stored to ensure proper handling of data.

Data is the most important asset in a business. We live in an era where companies find value in collecting and sharing data. The business had to meet legal responsibilities about the collection, storage, and process of personal data.

Data privacy issues and properly applying laws has increasingly contributed to the business for success. Perspective on objectivity and how they affect the applicability of various data protection and privacy laws have to be drawn. It creates a challenge to identify data controllers and data processors in various blockchain implementations. In distributed blockchain networks there is the territorial implication. A great variety of regulations can incur significant overhead costs. There should be potential restrictions when cross-border data transfer takes place. They require some centralized program to implement them. Difficult to implement in public blockchain with undefined groups. Applying criteria for processing personal data in the blockchain should be structured.

i. Data Threat Mitigation steps

Several risk management strategies can be developed when considering data privacy in blockchain technology:

- Use permissioned blockchain to support governance models:
    - Authorize selected number of approved participants.
    - Technical measures to reduce the amount of personal data that participants process.
    - Allocating data processing responsibly.
    - Responding to individual requests.
    - Deploying data processing agreement.
    - View differences between public and private blockchain implementation.
- Limit personal data stored in the blockchain:
    - Avoid putting personal data on a blockchain.
    - The financial system does not involve a naïve user.
    - Avoid payload for storing personal data on the blockchain.
    - Use one-time addresses to secure data in the blockchain.
    - Supply management chain to limit data on the blockchain.

ii. Blockchain Privacy Management

From the perspective of privacy compliance blockchain technology appear to be least ambiguous. Processing data on a public blockchain may involve significant business risks.

Suggestion from technologists:

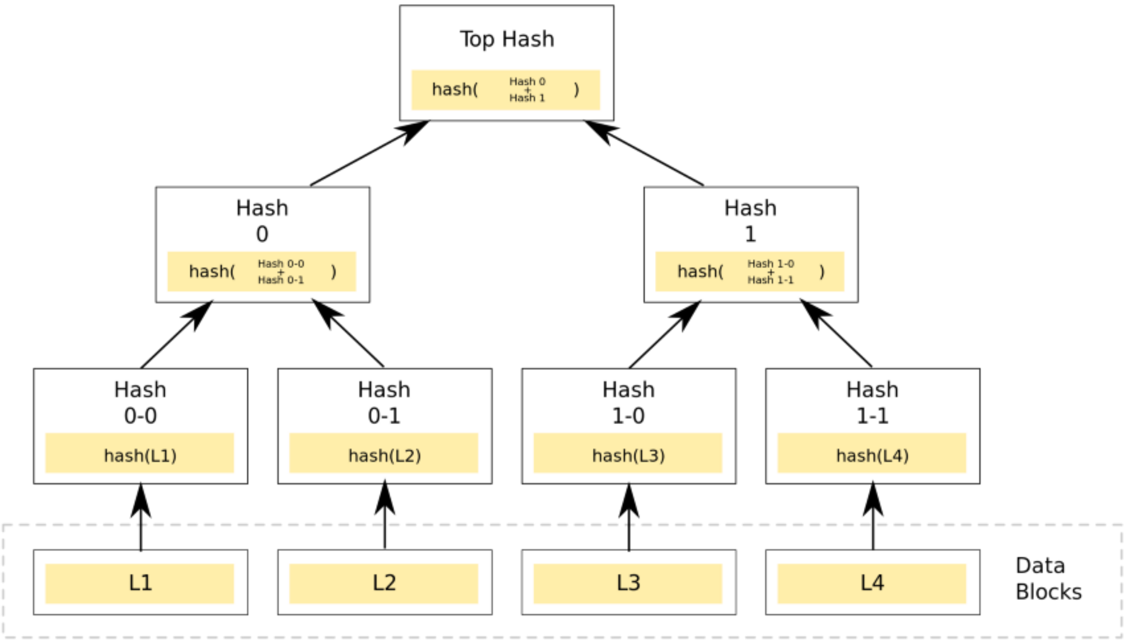
- Managing and verifying consent.
- Minimizing sharing of data between the data controller and data processor.
- Providing individuals with clear notification.

Self-governing blockchain-enabled identity and data management solutions to maintain data and privacy policy can be the potential solutions.

Payment Verification

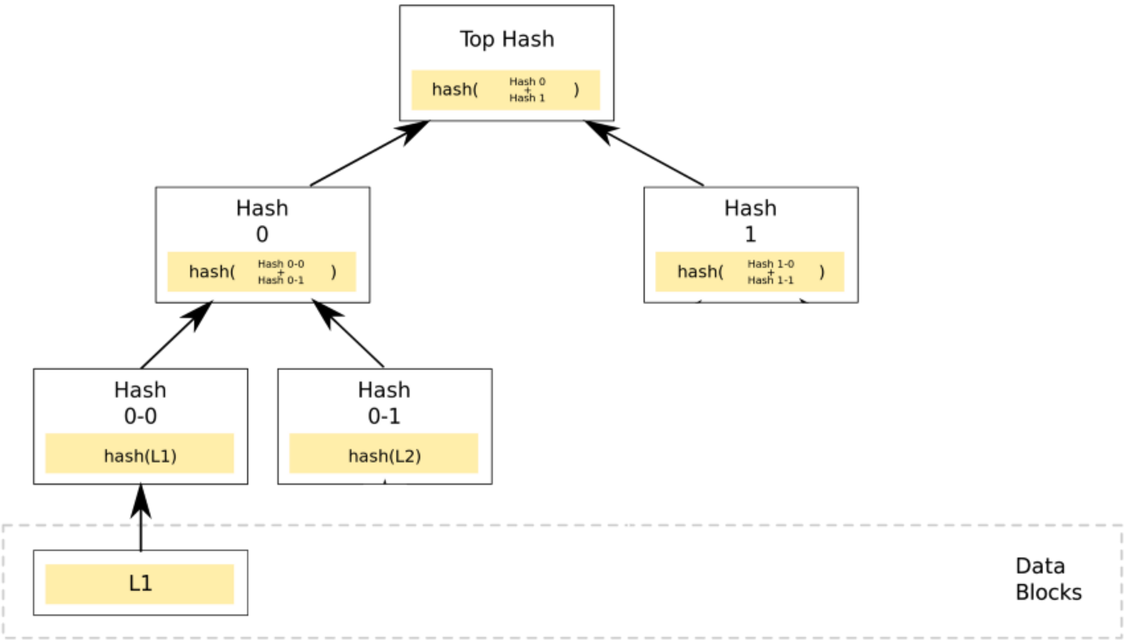
Simple Payment Verification, usually abbreviated to SPV, is a system that enables light clients (wallets running on low-end systems) to verify that a transaction has been included in Bitcoin and therefore a payment has been made.

This is possible by using the Merkle tree to store the transactions in each block. A Merkle tree is a structure created by grouping all the transactions in pairs and hashing them together, then proceeding to hash the resulting hashes together and continuing this process till there is only one hash left, called the merkle root. This creates a tree where every node has two children, which can be used to create their parent node.



Visualization of a merkle tree, L1-L4 are Bitcoin transactions

Someone that only knows the Merkle root/top hash can verify if a transaction is part of the tree, that is, if it's been included into a Bitcoin block. This is done by taking the nodes that are in the path that connects the Merkle root with one of the bottom transactions and bundling them together to create a proof:



A SPV proof constructed to prove that L1 is included in the block

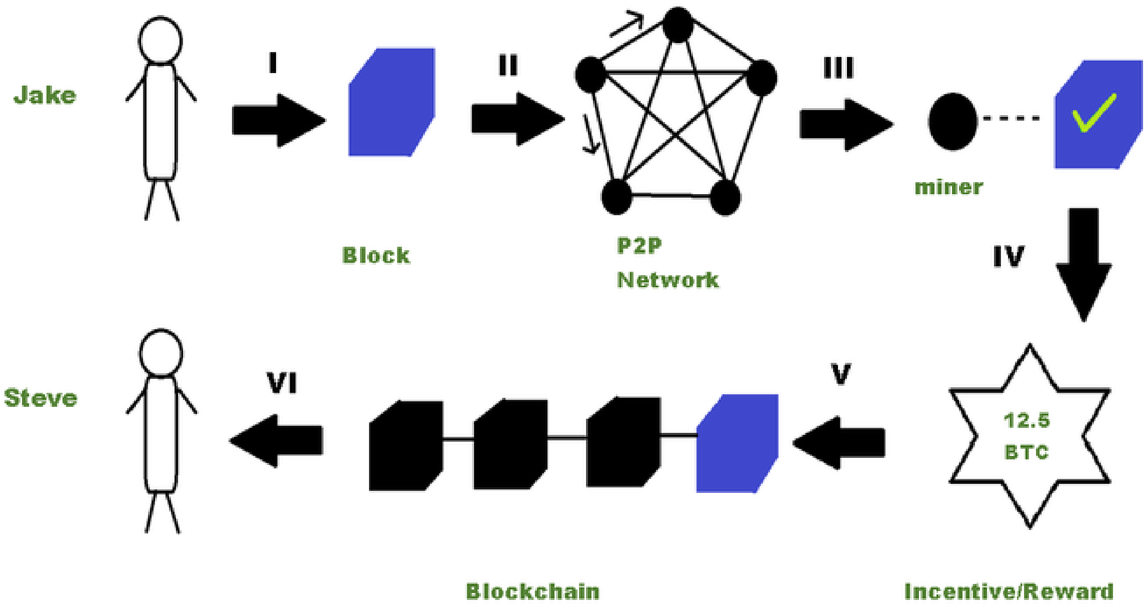
With that proof, our original user that only had access to the top hash can follow the path back to the roots in a verifiable way, he can check that Hash1 and Hash0 hashed together generate the top hash, meaning that Hash1 and Hash0 are its legitimate children, then apply this same check to Hash0–0 and Hash0–1, thus asserting that these two are also part of the original block, and, finally, check that L1 is the source of Hash0–0, proving that L1 is included in the block, therefore confirming it as an accepted Bitcoin transaction.

Running a full node requires downloading the entire blockchain, but if we use SPV proofs we only need to know the merkle root of each block in order to verify the transactions, so we only have to store 80 bytes per block, instead of the much larger size per block required for full nodes. This decrement of over 99.99% makes running the verification inside a low-resource device or a smart contract feasible.

Resolving Conflicts and Creation of Blocks

Bitcoin mining is the process of adding transaction records in the form of blocks to the blockchain and the nodes in the network that compete against each other to create a valid block and successfully add it to the blockchain are called miners. Since the process of creating a valid block requires a lot of processing power, there is an incentive mechanism, in which the miners are rewarded for their work via the Proof-of-Work (PoW) algorithm. The PoW algorithm is a consensus algorithm that is adopted in the blockchain network where the miners have to solve complex puzzles (which require a lot of processing power) and then they are paid in bitcoins for creating a valid block and successfully adding it to the blockchain.

The process of adding blocks (transactions) to the blockchain can be understood with the following diagram and its corresponding steps.

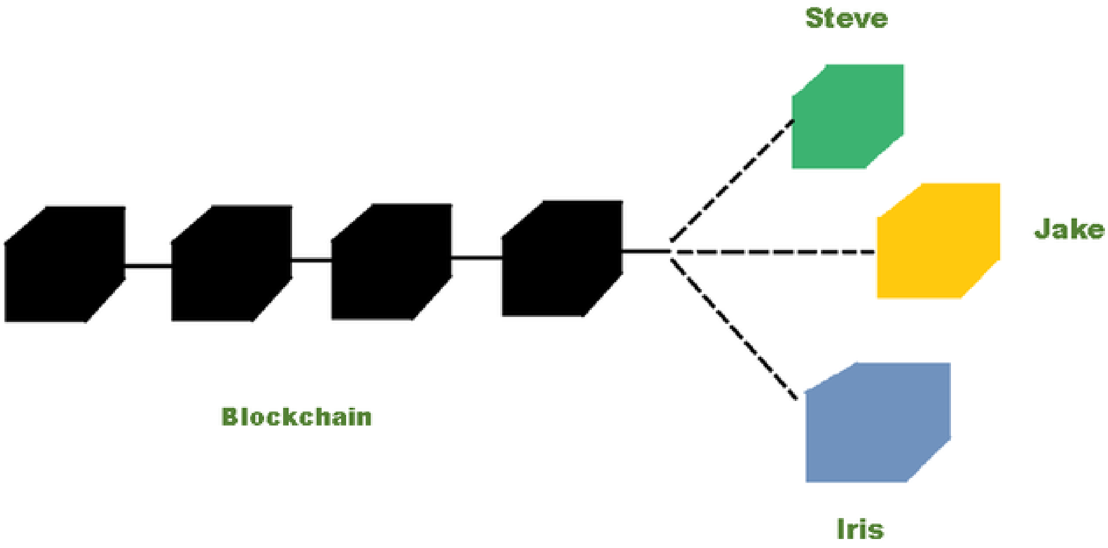


The steps involved in a Bitcoin Transaction are given below:

1. Jake initiates a transaction of say 15 BTC that needs to be transferred to Steve.
2. A block consisting of the transaction is flooded throughout the P2P network.
3. The miners validate the transaction via the proof-of-work consensus algorithm
4. An incentive is given for the miners who successfully create a valid block.
5. The new block consisting of the transaction gets added to the blockchain.
6. Steve receives the 15 BTC that was sent by Jake thereby completing the transaction.

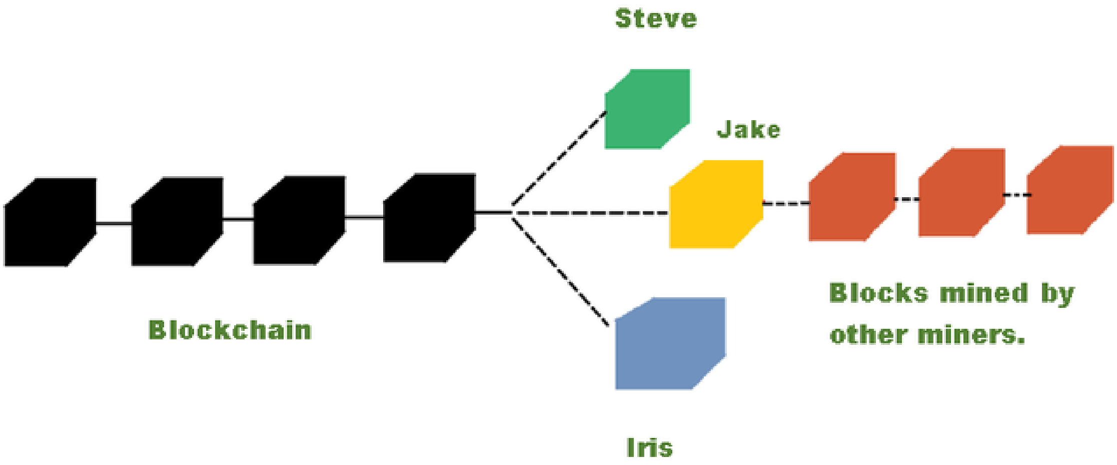
From the above-given steps, in step 5, a conflict can arise where multiple miners create blocks at the same time and try to add them to the last valid block of the blockchain. In this case which block will be appended to the blockchain?

Consider Steve, Jake and Iris are miners in the blockchain and they simultaneously create their respective blocks which are known as candidate blocks (represented in green, yellow, and blue respectively). Out of these candidate blocks, a decision has to be made to choose which of these blocks should be put in the chain.



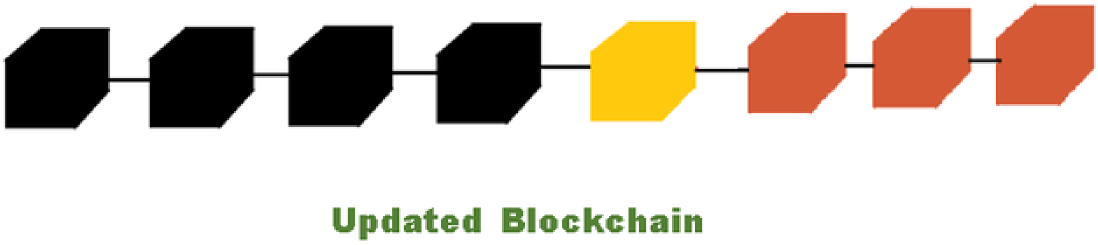
A conflict arises when Steve, Jake, and Iris mine their block at the same time

This conflict is resolved using **the longest chain rule** which is adopted by every node in the network to achieve consensus on the valid structure of the blockchain. To add a node in the blockchain, a miner in the network needs to have a computer with high processing power. In this case, let us say Jake's computer has more processing power as compared to the computers of Steve and Iris. So, Jake can create a block faster than Steve and Iris. Meanwhile, other miners are trying to create other valid blocks, so these blocks are created on the block which has occurred first, and since Jake's block was created faster than Steve's and Iris's, it occurs first in the chain and the miners add their blocks (given below in red) on top of Jake's blog (given below in yellow).



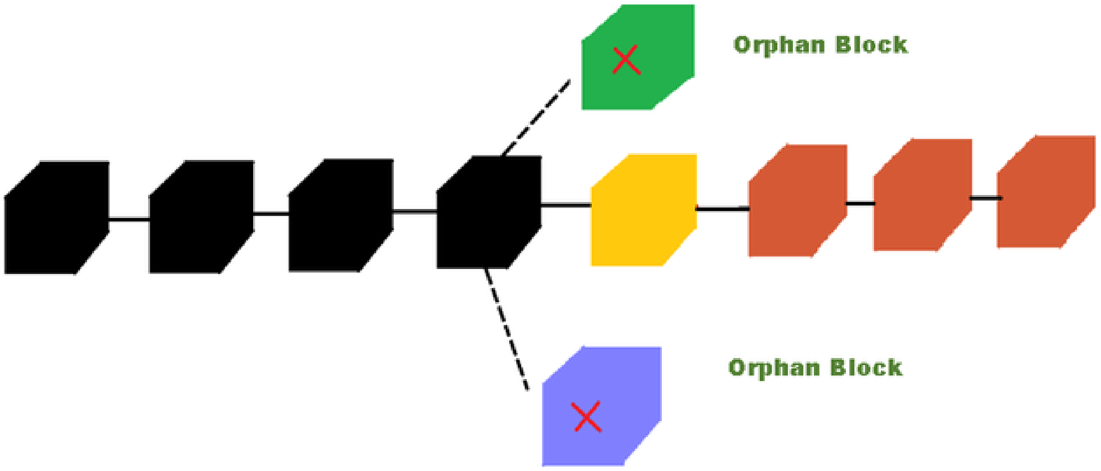
Other miners create their block at the top of Jake's block

It can be seen from the above diagram that the unit with Jake's block is the longest unit and according to the longest chain rule, the unit with the longest length must be accepted as the valid version of the blockchain. Therefore, the updated valid version of the blockchain looks like the one below:



The updated version of the blockchain is distributed in the entire P2P network

The blocks created by Steve and Iris are discarded and they become **orphan blocks** since they are not part of the main chain of blocks anymore.



Orphan blocks are discarded from the blockchain

It is important to note that the transactions inside these orphan blocks are valid because someone in the network has initiated it, but cannot be included in the blockchain as they do not follow the longest chain rule. So, these transactions are sent back to the **transaction pool** (which contains those transactions that are not yet mined), where the miners can pick them from the pool and then start mining them again.