**Blockchain tutorial**

Module A –

What is blockchain? – Hashing Algorithm – Immutable Ledger – Distributed P2P network – Mining? – Consensus protocol

Module B –

Bitcoin – Bitcoin´s monetary policy – Mining – Nonce – CPU, GPU, ASIC – Mempool – Transactions – Wallets – Public key and Private key

Module C –

Ethereum – Smart contract – Dapps – EVM and Gas – DAOs – Hard and Soft Fork – DAO attack – ICO – Ait coins

**Module A**

Hashing Algorithm –

* Block – block number, Nonce, timestamp, data(transactions), previous Hash, Hash.
* The Hash of the current block is generated by the SHA 256 algorithm.
* Any document or text or audio or video is put into the SHA256 algorithm to generate a hash value. The hash value is a 64-hexadecimal character. So, each character is of 4 bits. So, in total it has 64\*4 = 256 bits. That´s why it is called SHA256 algorithm.
* The first block in the blockchain is called Genesis block. The prev. hash of this block is all zeroes (00000000……0000).
* The five requirements of hash algorithm are - One-way, Deterministic, Fast computation, Withstand Collisions.

**One-Way** – It means the hash value is not reversible. Once the PT is converted into hash, the hash can´t be converted back to the PT.

**Deterministic** – The same PT generates the same hash value always.

**Fast computation** – The hash is generated very fast as is required to create the block.

**Withstand collisions** – It shouldn´t get hacked easily, otherwise it would be difficult to work with this hashing.

**Avalanche effect** – Even a small change like adding a space or a comma or a semicolon would completely change the hash value of the block.

Immutable Ledger

The Same blockchain is distributed among hundreds or thousands of P2P nodes. If in a blockchain, any of the block is tampered, the blocks succeeding it will all become corrupted. Now as soon as this blockchain doesn´t match with other blockchains, it will be discarded from the network.

Distributed Peer to Peer Network

The main disadvantage of a centralised network is that once it´s compromised, all the data that it contains can be tampered with, modified and used in any way possible. While, on the other hand, in the distributed peer-to-peer network, even if a blockchain is compromised and modified, the edited blockchain will be discarded due to not getting matched with the other blockchains in the network.

So even if a particular node adds a block in the blockchain or modifies blocks in the blockchain, it would be validated by the other peers in the network. The new blockchain will only be followed if that is accepted by the other majority of the nodes or else it will be discarded. So, we can say that it is very difficult to change the blockchains or corrupt it, and that is the main advantage of a distributed P2P network.

Consensus Protocol

Different types of consensus protocols are as follows – **Proof Of Work (POW)**, **Proof Of Stake (POS)**, **Delegated Proof Of Stake (DPOS)**, **Proof Of Authority (POA),** etc.

Diagram

Description automatically generated

In Consensus Protocol, always the longest block chain rule is followed. Here, for example, at the same time node A and node C has mined a block at the same time. Now Node E & Node F has validated the orange block mined by Node A and added to their respective blockchain. On the other hand, Node B has validated and added the purple block mined by Node C in its blockchain. So now, there are two different blockchains.

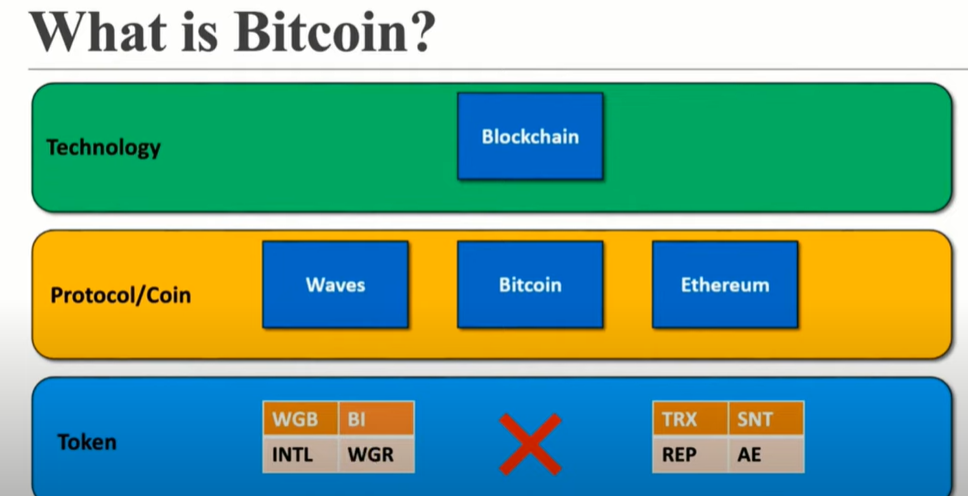
Diagram

Description automatically generated

In this scenario, node A, E & F will try to mine another block and similarly Node B & C will also try to do the same. Whoever mines a new block first will win the competition and that blockchain has to be accepted by the other nodes in the blockchain. Here, for example, we have considered that Node E has mined the sixth block and so after validated Node B and Node C has to adapt that blockchain only. The purple block will be discarded, and they are called **Orphan blocks**. Node C will not get any reward for its mining of the purple block and all transactions will be discarded.

Consensus Protocol is better than the Byzantine´s General solution as the later required 66% of the majority vote, while the former requires only 51% to be accepted.

**Module B**



Blockchain is a technology whereas waves, bitcoin and Ethereum are all protocols based on the technology blockchain. Tokens are developed based on the platforms like Waves, Ethereum, etc. Bitcoins have no token and they have no such things such as smart contract.

Blockchain´s Monetary Policy

Satoshi Nakamoto defined the monetary policy of blockchain majorly by two principles – **The Halving** and **the block frequency**.

Table

Description automatically generated

In the Halving policy, it said that after every 4 years or once the block number reaches 210,000 then the reward amount to the miners will be reduced by half. So, in 2009, the number of reward was 50 bitcoin and then once the number of blocks reach 210,000, it got reduced to 25 and so on. It is expected that in this rate, the reward will become zero in the year 2140. However, miners will continue to mine because there will be transaction fee awarded for every mining.

Block Frequency

It takes approximately 10 minutes to mine a block on an average. Some blocks may take 20 mins, some may take 8 secs but overall, on an average it takes about 10 mins to mine a block in a blockchain.

How mining Works?

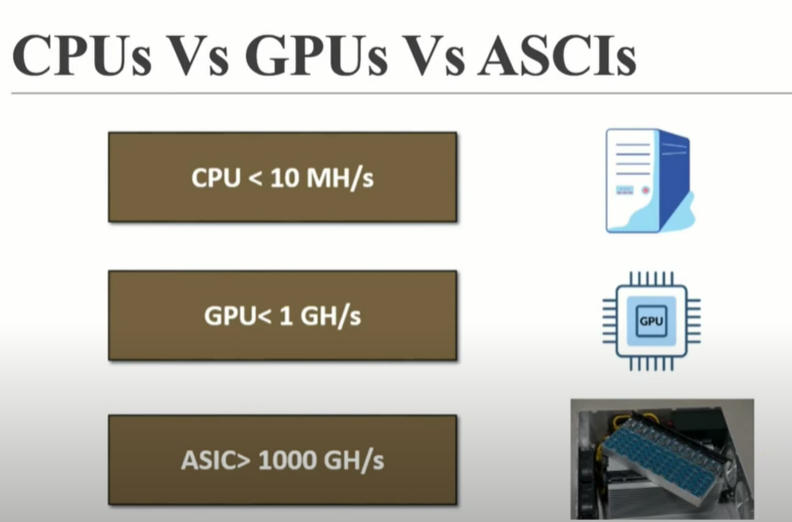
Graphical user interface, text, application

Description automatically generated

Miners need to solve a complex mathematical problem in order to generate a hash. Now, in a block all the fields remain the same and can´t be changed except Nonce. So, miners need to generate a particular nonce for which the hash value reaches the target set, only then the block and the transactions will be accepted and validated by other miners to be included in the blockchain.

The target is a value below which the hash needs to be generated by the miners. If the hash exceeds the target value, the block is not accepted as a valid block. So, by changing the nonce continuously the miners generate the hash and once a hash is generated that is below the target level, the block is accepted as a valid block in the blockchain.

Also, the blockchain algorithm sets the target for a particular network. The target usually keeps on changing after every 2016 blocks. Depending on the speed at which the network is mining the blocks, the target is set for that particular network. If a network is able to mine at a very fast rate and taking way less time than the average 10 mins, then the difficulty level of the target is kept high for that particular network and so on.



These are the different mining capabilities of different hardware. More the hashing speed less would be the time required for the miner to mine a particular block. Nowadays, ASIC has become powerful with approx. 14-15 Trillion Hashes/Sec.

Mining Pool

Diagram

Description automatically generated

* The issue with large companies with higher mining rate is that small miners can´t compete with them. So, if a small miner can join a group of other miners to form a mining pool then they can together compete with the larger mining players or companies.
* The reward is distributed among the members of a mining pool based on the hashing power an individual miner has utilised to mine that particular block.
* One more advantage is that the range of Nonce is divided into the miners so that no two miners search the same range of nonce. This makes the process fast with no duplication.

Nonce Range

The Nonce is a 32-bit number. So, Range of Nonce = (0 to 2^32)-1 = 4\*10^9. Approximately, 4 billion Nonce can be generated and resultantly 4 billion hashes can be generated.

Also, using SHA256 algorithm, we can produce 64 bits each out of a possible 16 characters. So total possible combination = 16^64 = 10^77.

Issue: Even though 10^77 hashes need to be checked, only 4 billion nonce can be generated by the miners. So, there is a possibility that the miner isn´t even checking the right hash that would make the block valid. To solve this issue only, the concept of timestamp comes into play.

Timestamp

The timestamp in the block is the Unix timestamp. It is running since Unix started.

Graphical user interface, application

Description automatically generated

To solve the issue of nonce, timestamp field has been introduced in the block. So, let´s understand the concept of timestamp.

* Suppose the miner can exhaust the 4 billion nonce in 40 secs.
* So, in 1 sec, he can exhaust 0.1 billion. So, in 0.5 seconds, he can exhaust 0.2 billion nonce.
* However, after 1 sec, the timestamp changes and so due to avalanche effect the hash also changes.
* Now the miner can reuse the Nonce he used in the previous timestamp to generate a different hash. In this manner, the Nonce value range will not be exhausted before hitting the right target value.

However, the current issue is not so simple as the hashing rate currently over a network is much higher than 0.1 billion/sec. It is approximately 180 million trillion hash/ sec.

Graphical user interface, text, application, email

Description automatically generated

This means the total range of nonce will be exhausted way before than even 1 second. So the question is: **what would the miners do in the idle time? Should they wait for the timestamp to change**?

Mempool

The above-mentioned issue is solved by the concept of Mempool.

Diagram

Description automatically generated

* In the Mempool, there are thousands of transactions. The miner initially picks up some of the transactions from the Mempool and puts it inside the block. Then, he performs the complex mathematical operation in order to generate the valid hash (hash below or equal to the target value).
* If the miner solves that within 1 sec and the timestamp hasn´t changed, then he picks the lowest fee transaction inside the block and replace it by the next lower fee transaction from inside the Mempool. Then again, he tries to solve the complex mathematical problem.
* This process continues until the timestamp has not been changed. Once the timestamp changes, then again he can bring the original set of transactions that he had chosen from Mempool.
* In this manner, the miner can reuse the transactions in the Mempool and the Nonce value range to find the valid hash.

Diagram

Description automatically generated

All the nodes and miners in the blockchain will have the Mempool. Now when a miner takes out the transactions from the Mempool and mine a block validly, then that info is passed to all other nodes and miners and the same result is reflected in all the Mempool and blockchain in the network.

Diagram

Description automatically generated

In our normal banking transactions, all the above transactions are added, and the total balance is determined and then it is decided whether we have the sufficient fund to buy the coffee.

Whereas, in the cryptocurrency market, no transaction is modified as we already know that it’s a distributed immutable ledger. So, the transaction from Alice (in this example) is considered to buy a coffee and the remaining 0.2 BTC is again listed as another transaction from (Me -> Me 0.2 BTC). No previous transactions is touched and only new transactions are added in the ledger.

Graphical user interface, text, application, chat or text message

Description automatically generated

UTXO – Unspent Transaction Output

In this example there is no Me -> Me transaction as we didn´t require the need for a change. And the first three transactions are used to pay for the noodles. So, after buying the noodles the only remaining UTXO will be “Bob -> Me 0.1 BTC”.

Diagram

Description automatically generated

Also due to the work and effort, miners put in for validating our transactions, we need to pay the miner a transaction fee. So, in this case we can see that a fee of 0.1 BTC is paid to the miner as a transaction fee and the remaining 0.1 BTC is only left with us. (Me -> Me 0.1 BTC)

Cryptocurrency Wallets

Diagram

Description automatically generated

Wallets adds up the transactions **that are not already spent** and **that are directed towards “Me”**. It then adds up the total balance and display the balance that I have in the wallet.

Public key and private key Cryptography

The transactions in the Mempool can also be done any hacker, which is not at all a valid transaction. And also as there is no central authority, there is no one to validate each and every transaction. So, to resolve this problem, the concept of public & private key cryptography comes into play.

Diagram

Description automatically generated

Every transaction done by the nodes can be considered as the messages here. Every node also has his/her own private key, which is used to generated his/her public key. Now the transactions(here messages) are encrypted by the private key to generate a digital signature. The Message, the Signature and the Public Key are put into the verification function, which the miner uses to check the authenticity of the person doing the transaction and the transaction itself. If the Verification function returns Yes/No, then it is a valid transaction or else there is fraud.

Segregated Witness

Diagram

Description automatically generated

* The Size of the blockchain was decided to be 1 mb. However, around 2017, it became a problem as there were thousands and thousands of transactions incoming and it was getting difficult to accommodate many transactions in a block.
* Also, the signature and public key within a transaction ID itself occupies 60-65% of the total size.
* That´s why the blockchain community decided to send the public/private key separately so that more and more transactions can be occupied inside a block. This concept is known as Segregated Witness.

Bitcoin Address

Diagram

Description automatically generated

The concept of Bitcoin Address was introduced to add an extra layer of protection. Even if some hacks our bitcoin address or tampers with it, maximum he can reach is to the public key and not to the private key.

**Public Key** – It is only used when we want to send a transaction to someone.

**Bitcoin Address** – It is used when someone is sending us some bitcoin. We will provide that person our bitcoin address so that we can receive the transaction.

Hierarchically Deterministic (HD) Wallets

Diagram, shape

Description automatically generated

To introduce an extra layer of security, the concept of Master Private Key is introduced. Master Private Key is used to generate Private Keys. Even if 1 is added, just like avalanche effect, it would generate a completely different Private Key. In this manner, with the help of Master Private Key, different private keys can be generated, which in turn can generate different public Keys.

Also, on the other hand, there is a Master Public Key. But the person who has the Master public key can only check the activities done by the other public key holder. He/she doesn´t have the access to the Master Private Key and hence will not be able to tamper with the private key – public key combination or any transactions.

**Module C**

What is Ethereum?

Open source blockchain based platform. The advantage over bitcoin is that it offers its own token which bitcoin doesn´t.

The concept of running programs on blocks and not only transactions gave birth to the idea of Ethereum. Consequently, came the idea of smart contract, which is usually written in the language of solidity on Ethereum platform.

Nodes of Ethereum

There are three types of Ethereum nodes – Full node, Light node, Archive Node.

**Full Node** – Locally stores a copy of the entire blockchain. Verifies and validates all the blocks.

**Light Node** – stores only the block header. Depends on full node. For low-capacity devices, which cannot afford to store the gigabytes of data.

**Archive Node –** Stores everything kept in the full node and built an archive of historical data, starting from the genesis block. It requires terabytes of diskspace.

Accounts in Ethereum

An Ethereum account is an entity with an ether (ETH) balance that can send or receive transactions on Ethereum. There are 2 types of Ethereum accounts –

**Externally Owned Accounts (EOA)** and **Contract Account (CA)**

**EOA -**

**Diagram

Description automatically generated**

The EOA account is automatically created when we do a transaction on the Ethereum platform. We need a Wallet and a Private Key to perform several operations such as send, receive transactions and also interacting through smart contracts.

**Contract Account (CA) –**

It is controlled by the Contract Code.

Diagram

Description automatically generated

The Contract Account is controlled by Smart Contract. In this example, if the delivery is done from A to B maintaining the right temperature, automatically the money is debited from B and paid to A. None of the parties can change the contract once written because of its immutable property and the contract will execute once deployed. Whenever the smart contract is deployed on the Ethereum block, the Contract Account (CA) is created.

**Differences between EOA and CA**

**Table

Description automatically generated**

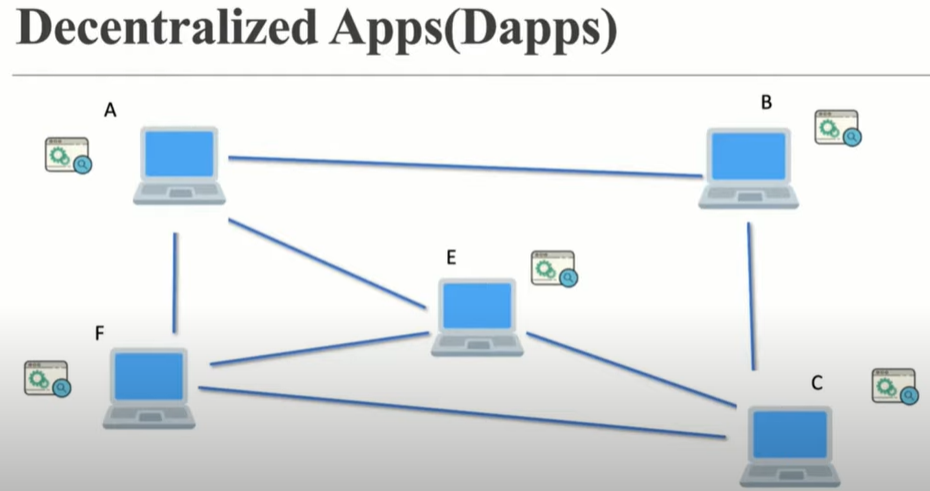
Both EOA and CA holds a unique address so that the involved parties can access the account and make transactions or view smart contract. Both holds ETH (ether) balance. However, in EOA we don’t need to pay any Gas amount whereas in CA we need to pay a certain Gas amount for execution.

Smart Contract

* Smart Contract is a program, usually written in Solidity language, that runs on the Ethereum platform.
* The language used in bitcoin is Bitcoin Script, which is Non-Turing language, while the language used in Ethereum is Solidity, which is a Turing language. Turing means that whatever we can think of can be written in the code and the code will run on a loop.
* Bitcoin doesn’t have the feature to run code on a loop as because the community doesn’t want any hacker to intervene and run a piece of code infinitely and slows down the network.
* However, in Solidity, there is a payment involved to execute the contract deployed on the Ethereum platform. So, every execution will cost a certain Gas amount from the parties involved to run the contract.
* Each node in the network has the following – 1) Current State of all Smart contracts and 2) History of both transaction and smart contract.

Decentralised Application (Dapps)

In a centralised system, there is 1 server and all the clients request to the central server if they require anything. However, the concept in Dapps is very different.



The application runs in all the nodes in the network and is not centralised to any particular node. That´s why it is called Decentralized Apps.

Dapps is created by two things – Smart contract (acting as the back end) and the Front end (for the human interactive part). Using these two components, we can deploy the app in the whole decentralized network.

Some examples of Dapps are –

Graphical user interface

Description automatically generated

Google search engine, twitter and you tube are all web 2 applications and run on a centralized server. On the other hand, similar types of apps running on Web 3 are Presearch, LBRY and Dtube. Some advantages of Dapps over Centralised applications are as follows -

* Centralised apps are not trustworthy, normal public can´t really check all the conditions and algorithms they are following, whereas in decentralized apps, anyone can check whether the conditions are getting fulfilled or not. So, they are trustworthy.
* For watching advertisement on google and YouTube, we pay, whereas in Dapps they pay the audience in crypto to watch the ads.
* Central servers, if go down, can halt the service for a certain period of time. There can be interruption of service of billions of users. Whereas the service can’t go down in Dapps. Even if the application or service is running in 1 node, it will continue to work, and it can’t go down.

Ethereum Virtual Machine (EVM)

EVM is a software which runs inside our Operation System and by the help of which we run the Ethereum Platform and can do all sorts of operations. That’s why the Ethereum software running on EVM can´t even touch the data in the local machine, i.e., the hard drive or the webcam of the local system. This actually reduces the risk of personal info getting stolen by hackers even if they the Ethereum software as that program (virus/malware) can’t access the local hard drive of the user.

Ethereum Gas

Every operation that is required to be executed in the smart contract needs a specific amount of gas, which is already set in the community.

Gas is equivalent to the currency which the sender needs to pay in order to execute the smart contract (amount the sender wants to pay per unit of the gas). The sender can set the price of the Gas that he wants to pay for a particular transaction. The price of the Gas is usually set is gwei. (1 gwei = 10^-9 ETH).

So the question is why would the sender pay a higher amount of gas if the work can be done in a lower amount? – The answer is because higher the payment of gas, faster the transaction is mined, and the operation is completed.

Ethereum Gas Limit

It is the maximum gas that the transaction can consume. This gas limit is set by the sender.

**Question:** Suppose A wants to send B 2 ETH. Then what will be the total fees that A has to pay for this particular transaction?

**Answer:** Let´s divide the answer into 3 cases.

*Case 1*: When A sets the transaction gas limit as 21000, which is the standard limit for a transaction.

So, if A sets the price per unit as 100 gwei. Then the total price = 100\*21000 = 2100000 gwei, which is equal to 0.0021 ETH.

*Case 2*: When A sets the transaction limit as 20,000 which is less than the standard limit of 21,000.

The transaction will fail. Moreover, A will not get the refund of the paid amount because the reward has to be paid to the miner, who has no idea of the set transaction gas limit by the sender. So, the loss is totally on A.

*Case 3:* When A sets the transaction limit as 22,000 which is more than the standard limit of 21,000.

The transaction will be successful. Also, A will get refund the price of the unused 1000 transaction gas.

**Question:** What is the use of the Gas limit?

**Answer:** There are 2 uses of the Gas limit – 1/ to set the transaction limit, 2/ to protect the user from DOS attack or being victimised from extra money being deducted from wallet due to infinite loop. If the transaction is consuming more than the set 21,000, then the program is instructed to stop immediately. That is a benefit of the Gas limit, which is set by the Sender at the very beginning.

Decentralized Autonomous Organization (DAO)

It is completely an autonomous organization run by the smart contracts in different nodes. There is no hierarchy and the complete process is totally transparent for the public. Main differences between the DAO and a traditional organization are listed below -

Table

Description automatically generated

The DAO Attack

The DAO is a real organization that works on the basis of a smart contract on the Ethereum blockchain Platform.

A picture containing diagram

Description automatically generated

The DAO basically functions as an investment firm. Venture capitalists have to provide Ether to the DAO, in exchange for which they will be provided with DAO tokens. With the help of these DAO tokens, they can either accept or reject in the investment proposals made by the DAO.

Timeline

Description automatically generated

There was a loophole in the smart contract written by the DAO due to which some hackers were able to withdraw 60-65 million dollars. So, basically it was a hit to the huge investments that were already made. The first proposal to counter the attack was given within a day and finally to stop such instances from happening again, the concept of Hard fork was deployed after almost a month.

After hard fork was deployed, we usually have 2 types of Ethereum coins – Ether & Ethereum Classic.

Hard Fork

During a hard fork, software implementing a protocol and it´s mining procedures is upgraded. Once a user upgrades it´s software, that version rejects all transactions from older software, effectively creating a new branch of the blockchain. However, those who retain the old software continue to process transactions. In a nutshell, two different communities run two different types of software and follow different blockchains after forking.

Bitcoin is also hard forked on July 2017 and bitcoin cash (accepting a new block size of up to 8 MB) was introduced. Demonstration of Hard Fork through an example -

Chart, box and whisker chart

Description automatically generated

Soft Fork

Soft Fork are changes to a protocol, but the end product remains unchanged. A Soft fork is a backward-compatible upgrade, meaning that the upgraded nodes can still communicate with the non-upgraded nodes.

Old nodes (not upgraded yet) can still validate the newly mined nodes but couldn´t understand them.

Chart

Description automatically generated

Also, ultimately the old nodes have to upgrade their software to join the majority of the new nodes, because anyway as per protocol the longest chain rule is going to be followed. So time and again, when a block will be mined by any of the nodes following the old protocols, tat block will eventually be discarded because of the longest chain rule.

Initial Coin Offering (ICO)

Diagram

Description automatically generated

Above picture is of IPO. In this format, public can invest their own money in return for which they purchase shares of the company. Also, in this format there are a lot of legal works involved. In addition to it, the more number of shares people buy, the more authority/Power they can have over the company.

Graphical user interface, application

Description automatically generated

Above picture is of ICO. In this format, public can invest cash in the form of bitcoin or ether and they can get tokens of the particular listed coin in return. Also, in this format, unlike IPO there aren´t that much of hectic legal work is involved. In addition to it, irrespective of number of shares, public can never have authority over the protocols of the coin or coin itself.

ETH 2.0 or Serenity

The major advantages of ETH 2.0 over the older ETH are as follows –

Scalability, Security and Sustainability.

The major upgrades that ETH 2.0 is having are Proof of Stake (PoS) and Sharding. Because of these reasons, ETH 2.0 is more acceptable and popular in the blockchain community.

Proof of Stake

Unlike Proof of Work, in this process, there is not much competition between the miners to become the first one to successfully mine and validate the blocks. Here, instead of miner, the system randomly selects the validator. More ETH someone pays as the stack (32 ETH minimum), higher is the chance of becoming the validator, being randomly chosen as the validator.

A picture containing text, clock

Description automatically generated

Once the validator successfully mines the blocks, the number of ETH he puts as stake is returned to him and he also gets a transaction fee as a reward for all the transactions being validated and mined into a valid block. On the other hand, if the validator tries to perform some malicious activity, the stake and the transaction fee is not given back to the validator.

The differences between the Proof of Work (PoW) and Proof of Stack (PoS) are as follows-

Table

Description automatically generated

Sharding

There are many concerns over the methods of Proof of Work –

1. The miners get reward only for mining. In cases where miners do try to mine the block in the first place and fail i.e., unable to mine the block at first place in the network waste energy in the process of doing so.
2. Even though only 1 miner is able to mine the block in the first place, all the other miners have to validate it before accepting it in the blockchain. This again is a loss of huge resources and electricity.
3. Moreover, majority of the time of miners is getting wasted in validating the work of other miner, instead of trying to mine for which they get the reward.
4. Also, in reality there are thousands of miners in a network, and it is again very important for every miner to maintain a long blockchain.

Due to these points, the concept of Sharding comes into picture. In Sharding, the whole blockchain is divided into smaller databases and each small database is being validated by a particular section of the network. The Other section of the same network validates the other section and so on.

So, the main advantages of Sharding are –

1. Transactions per second increases, because miners don´t invest a lot of time in validating as they need to verify only a smaller portion of the database. So instead, they can focus on mining and hence transactions also increase.
2. Powerful and expensive computers are not required as because they don´t need to validate or maintain the whole blockchain.
3. More validators will join as because powerful computers are not required and also anyone can become a validator only by putting 32 ETH in stake.
4. Energy consumption will reduce as because miners don’t need to maintain or validate the full blockchain over the whole network.

In the near future, many changes will deploy over the Ethereum blockchain bring many more advantages.

Diagram

Description automatically generated

Beacon chain will be deployed on Ethereum Mainnet and then it will be integrated with the Shard Chain. Maybe, people will use the bitcoin less because of it´s disadvantages over the Ethereum and its upcoming changes if the former continues to run on the Proof of Work Consensus protocol.

Altcoins

Any coins other than bitcoins are termed as Altcoins.

As of 2021, Ethereum and Binance coin are the most popular Altcoins with respect to the market capitalization.