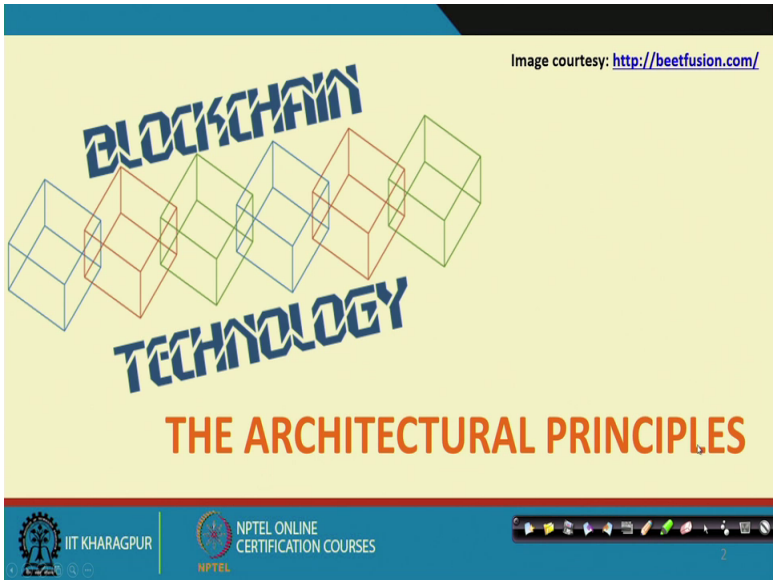
Blockchains Architecture, Design and Use Cases
Prof. Sandip Chakraborty
Department of Computer Science and Engineering
Indian Institute of Technology, Kharagpur

Lecture – 03
Introduction to Blockchain – III
(Architecture Principles of Blockchain)

Welcome to the course on Blockchain. So, this is our third class where we will discuss about the architectural principles behind blockchain.

(Refer Slide Time: 00:29)

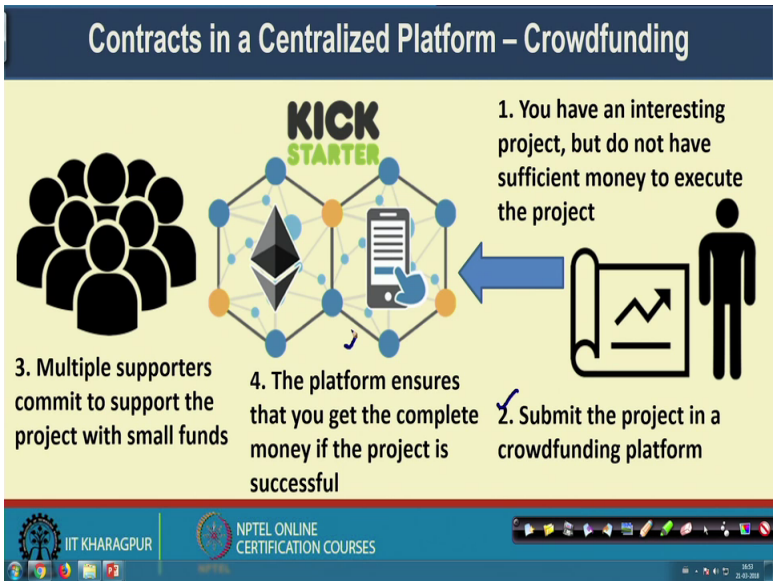


(Refer Slide Time: 00:32)

So, in the last class we have started discussing about smart contracts. So, this smart contract the term it was first coined by Nick Szabo who was a computer scientist and cryptographer. So, in 1996 he had first coined this term.

So, Szabo claimed that this smart contracts that whenever you are establishing some kind of contracts between multiple parties that can be realized with the help of a public ledger. So, that was his idea and because it can be realized by a public ledger the same concept can also be realized using a blockchain. So, blockchain can be a pioneering technology to realize smart contracts.

(Refer Slide Time: 01:17)

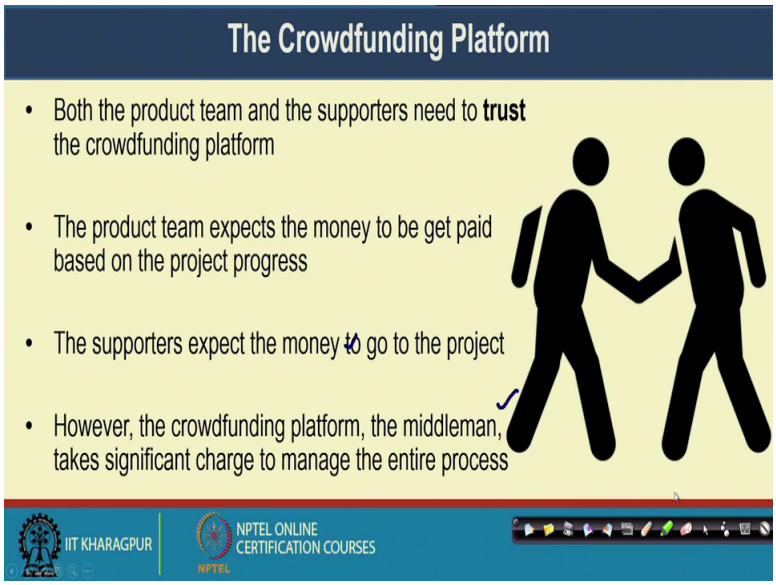


So, let us look into this concept of contracts in a centralized platform with an example of crowd funding. So, so there is some kind of crowd funding companies like Kickstarter. So, Kickstarter works in this principle like a say you want to execute some kind of interesting project, but you do not have sufficient money and this can be an individual or a group of people who want to execute some project, but they do not have sufficient money. So, what they do? That they submit this project to Kickstarter and means Kickstarter type crowdfunding company. And now there are multiple supporters who can support with some small funds to that particular project.

Now, the project can be executed by one or individual who have submitted the project proposal to Kickstarter. And on the other hand you can have multiple supporters; so, this multiple supporters they can everyone can support with a small fund and the total fund that you are getting with that help of that fund you can support that project or you can execute that project.

Now, the task of this Kickstarter platform is to ensure that when you are completing some milestone of the project, you are getting that fund. So, Kickstarter is ensuring that whenever some supporters is providing with the fund; the fund is going to the intended project and as an when the project completes some milestone. So, the project executers they are getting the fund and if the project is not completed successfully or in between the project gets scrap then the fund is sent back to the supporters.

(Refer Slide Time: 03:11)

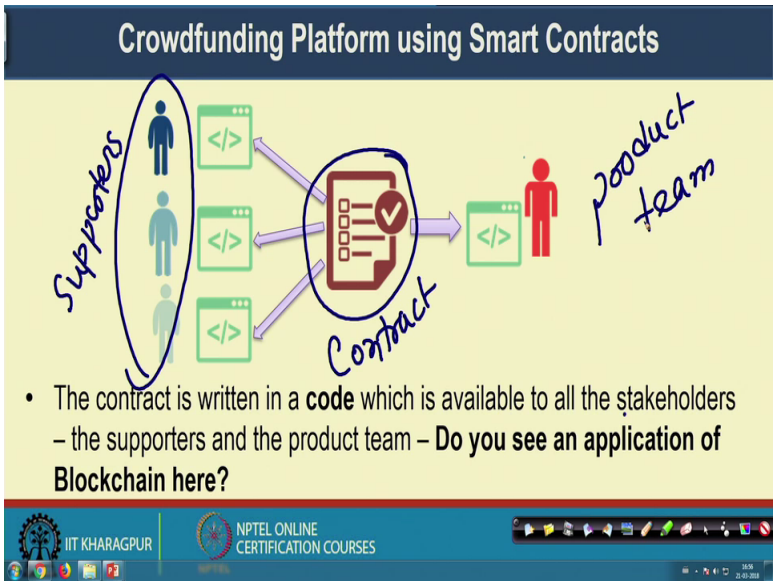


Now, in this kind of architecture you need to have a kind of trust relationship; that means, the product team they expect that their money to get paid based on the project progress. So, whenever there is some milestone that has been reached. So, they will be get paid by Kickstarter; the supporter they expect that their money is going to the right project and if the project get scrapped in between then they will get back their money.

Now, this crowdfunding platform; that means, the Kickstarter who is working here as a middleman they take significant amount of charges both from the supporters as well as a from the product team. So, that way a huge money is taken by the middleman and that is indeed the kind of problem with the centralized platform; that first of all you need to trust the platform and second that you have to provide a significant amount of charge to this to this middleman's because they are actually handling the risk factor.

So, you can understand for that for this kind of project there is a risk factor which is associated to it and because they are this kind of Kickstarter of the middle mans, they are handling this kind of risk factor that the project might not get completed or in between the supporters may claim that I do not want to support it further. So, this kind of risk is associated to it. So, the middleman like kickstarted; they are taking significant charge to handle this kind of risks.

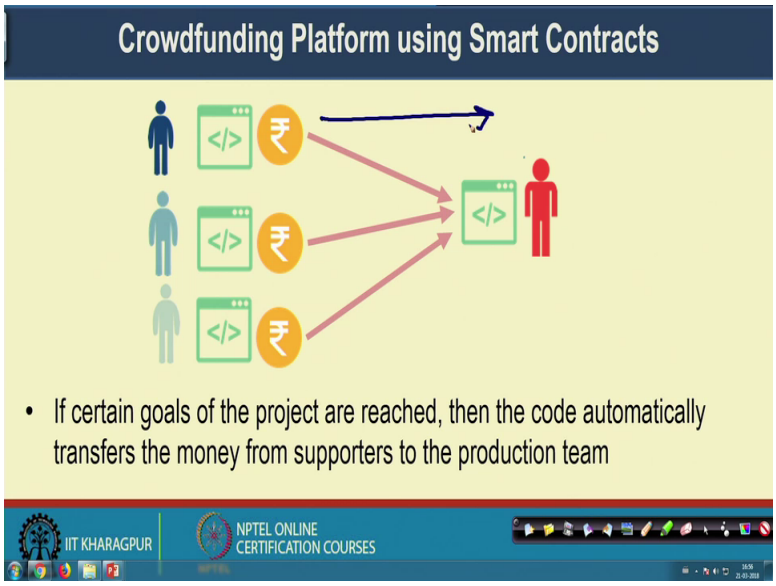
(Refer Slide Time: 04:44)



Now, let us see that how this kind of crowdfunding platform can be realized with the help of a smart contract. So, here you have a set of supporters; so, this are your list of supporters and on the other hand you have the product team.

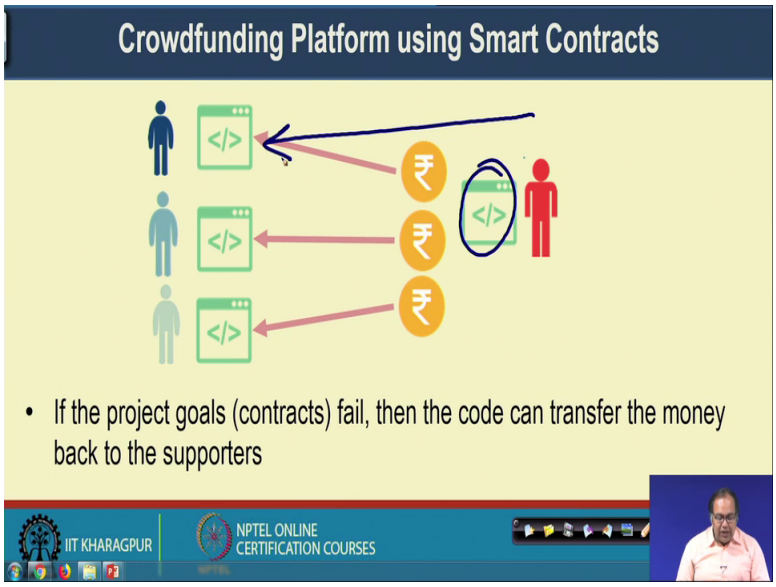
Now, this contract between the supporters and the product team it is written in a code which is available to all the stakeholders like that particular code. So, this code contains the contract; so, this is your contract and contract is made available to the supporters, as well as to the product team. Now everyone can verify that contract and in this particular case you can see that if we put that contract inside the blockchain; then everyone will be able to verify that contract, but no one will be able to tamper with that contract. So, this gives an interesting idea that well such kind of smart contract platform can be realized with the help of a blockchain.

(Refer Slide Time: 05:50)



So, whenever this contract get executed. So, the contract has like after 10 days if this particular project milestone has reached then you transfer the money from the supporter to the product team. So, if this condition get satisfied then the money is transferred from the supporter to the product team. So, that is automatically based on whatever is written inside the contract.

(Refer Slide Time: 06:18)

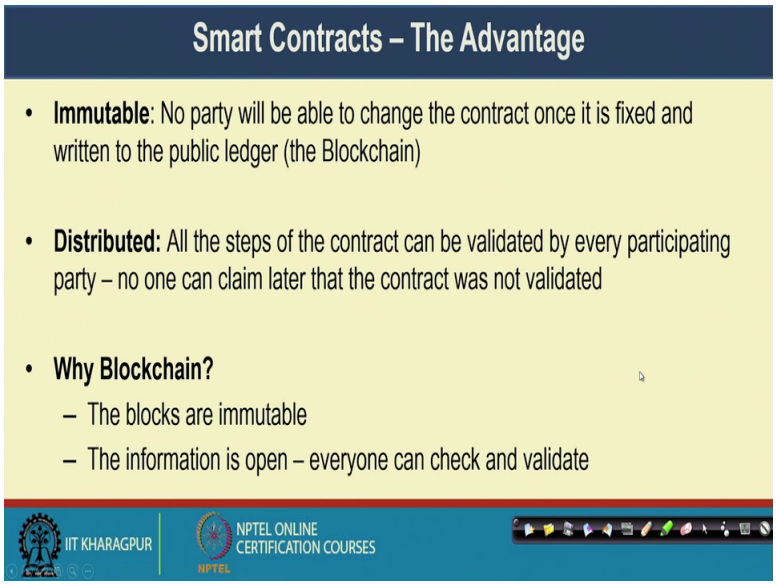


Now, whenever the project gets scrap if it is like that after some 50 days the product team says that well we are not able to make sufficient progress in the product. So, we

want to scrap the project then the money goes back from the product team to the supporters. So, the contract the code which is written there inside the smart contract; it automatically transfer the money from the from the product team to the supporter.

So, that way in this particular application inside the blockchain rather than putting some transactions or putting some data, we are putting a code which will be automatically verified by every stakeholders, they will not be able to tamper the code, they will not be able to deny the code in between. But as an when the code runs by verifying that whatever actions or whatever events have been executed the contract can get executed over time and fulfill the initial agreements that have been made.

(Refer Slide Time: 07:24)



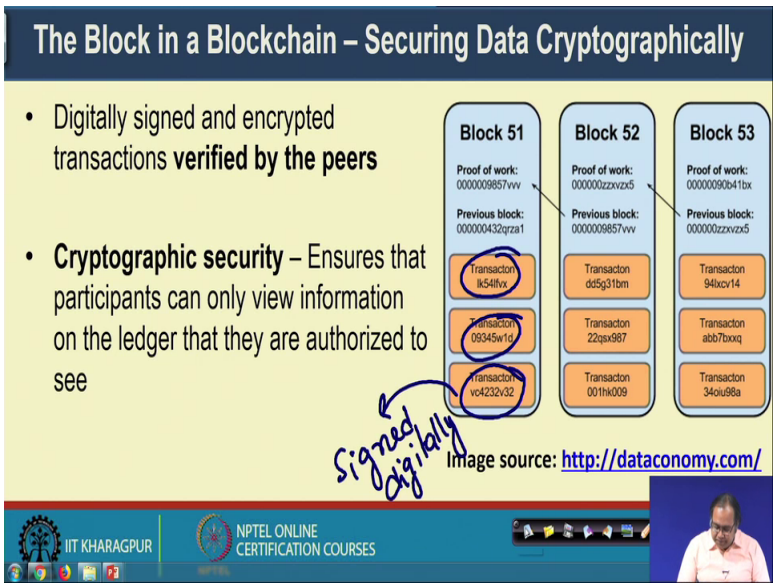
So, the advantage of this kind of smart contract is first of all they are immutable; that means, no party will be able to change the contract once it is fixed and return to the public ledger like blockchain. Next it is distributed you do not need a middleman like this like this say Kickstarter who is handling all the risk.

Like your code will be automatically get executed and if you are not fulfilling your promise then automatically the code will execute some steps based on the contracts. And why the blockchain? Why blockchain is a suitable platform for executable smart contract? Because first based on the blockchain architecture, the blocks are immutable and the second the information is open, everyone can check and validate the information inside a blockchain.

(Refer Slide Time: 08:17)

So, there are multiple smart contract platforms like ethereum, hyperledger, rootstock, ripple. So, we look into the hyperledger in detail; so, Praveen in his lecture he will explain the hyperledger platform and how you can write smart contracts using hyperledger.

(Refer Slide Time: 08:34)

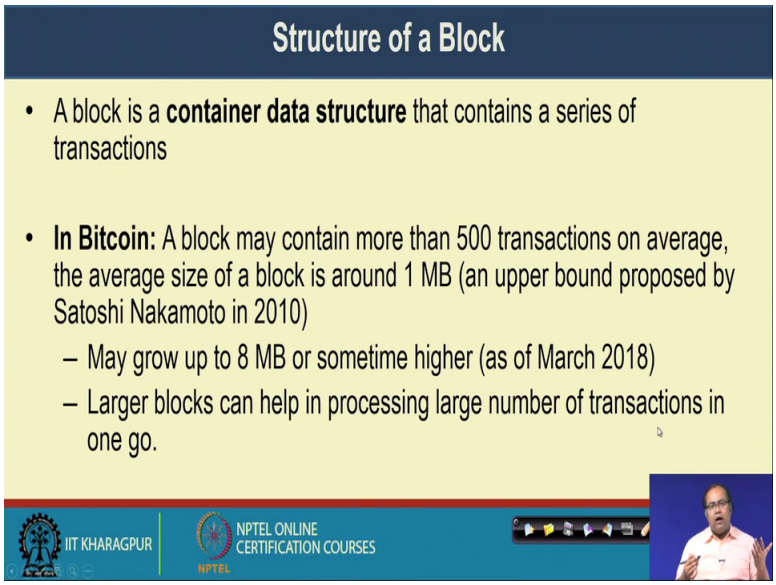


Well now, let us look into the blockchain in details little details. So, first we will look into what is there inside the block of a blockchain. So, we so to put some data in the blockchain; we want to secure that data and how a block is securing the data by utilizing

the concept of cryptography. So, we will take the example of bitcoin in this case to explain you that what is there inside the block and how individual blocks are getting connected. So, this blocks they are kind of they are containing the digital is signed and encrypted transactions which are already verified by the peers.

So, you in inside the block you can have multiple transactions. So, these are the transactions which are verified by the peers and this transactions are there in a encrypted format or it is it is basically signed digitally. These are digitally signed transactions which ensures that the participants they can only view the information on the ledger that they are authorized to see.

(Refer Slide Time: 09:47)

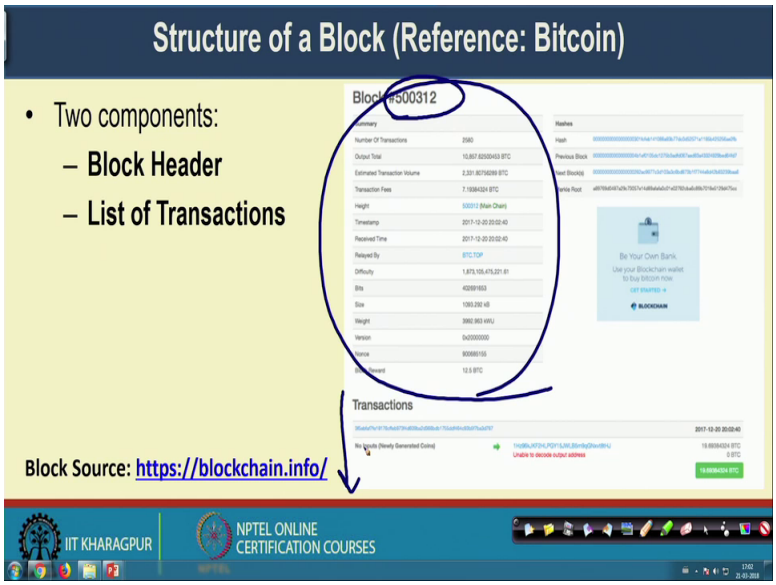


Now, if we look into the structure of a block a block is a container data structure that contains a series of transactions. So, in case of bitcoin a block may contain more than a 500 transactions on average. So, the average size of a block is around 1 MB; so, in the original white paper by Satoshi Nakamoto and bitcoin he has me he had mentioned that the upper bound of a block can be around 1 megabyte, but in now a days we are expanding the block size.

So, a block may grow up to 8 megabyte or sometime even higher than that. So, that was the information as of March 2018; the recent information and the larger blocks it can help in processing large number of transactions in one go. So, if you remember the mining procedure in bitcoin that the miners collects all the transaction together put them

in a single block. So, if you can put more transactions in a single block then you can process it in one go. So, that is the advantage of having a larger block, but there are multiple disadvantages that will discuss while we will talk about the consensus mechanism.

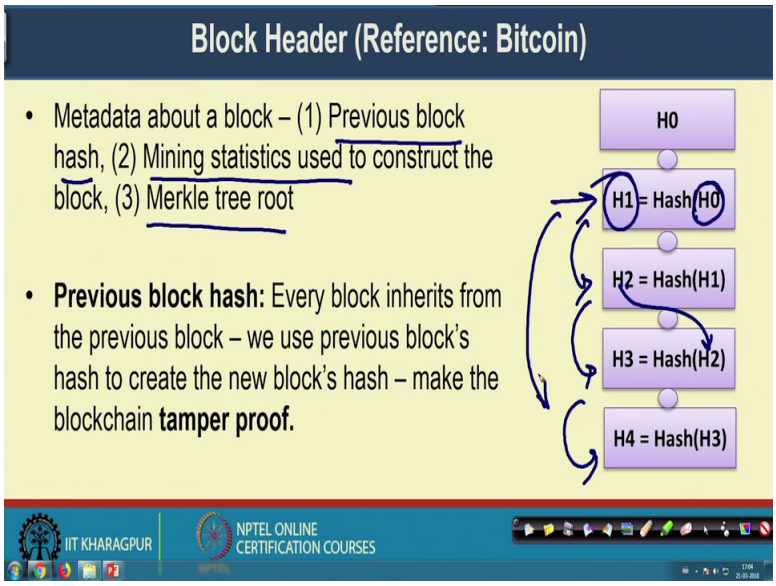
(Refer Slide Time: 10:51)



So, now let us look into the structure of a block; so, a structure of a block in the context of bitcoin they have two component the block header and the list of transaction. So, this is an example that is taken from this website blockchain dot info. So, if you go to that website you can see the current blockchain which is utilized to realize the bitcoin money transfer.

So, whatever block is there in the bitcoin network you can see all the block blocks by going to this blockchain dot info website. So, this is the information of a particular block with block number 500312. So, there are two part of the block the first part is the block header; so this is the block header and then there a list of transaction size of just shown one transaction, but there are list of transactions which are there inside that block.

(Refer Slide Time: 11:45)



Now, in the block header; so this blockchain as you know as I have mentioned that this blockchain is a sequence of blocks which are connected by the hash of the previous block. So, H 1 is connected with the hash of previous H 0, H 2 is connected with the hash of H 1 then this H 2 is used to connect H 3. So, that way the hash function actually construct the chain kind of structure.

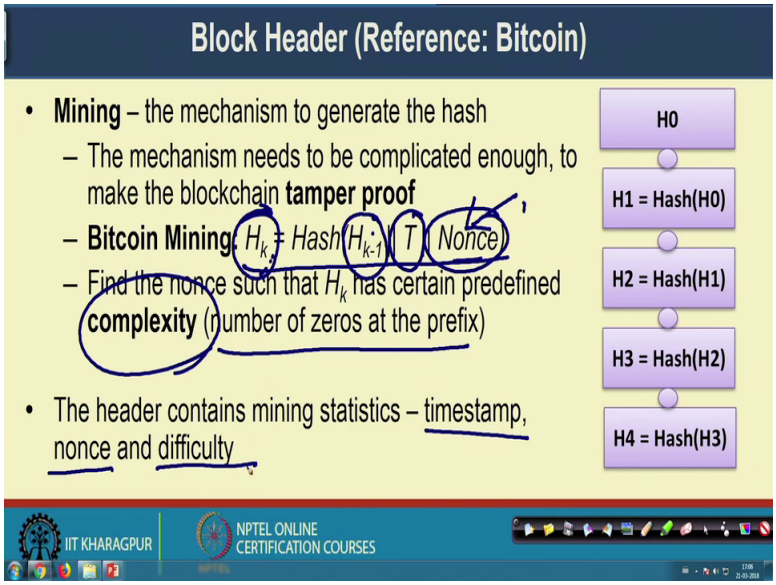
So, inside a block header you have this previous block hash which is utilized to construct the current block hash. So, here this is the previous block hash which is used to construct the current block hash; then the mining statistics in case of bitcoin that some statistics about mining I will come to that point. And something called the Merkle tree root which construct, which stores the information or a or a hash value of all the transactions which are there.

Now, this previous block hash; that means, in case of blockchain as we have mentioned earlier that every block inherits the from the previous block; that means, we use the previous block hash to create the new blocks it makes the blockchain tamper proof; that means, if you want to make some changes saying in this block; that means, this particular hash value will get changed and you have to change all the subsequent hash values.

So, that way you can think of that in a distributed network some people is trying to tamper the block if that person is trying to tamper the block; he has to make change in all the blocks which are there after that. And we want to make this problem as complicated

such that by the time some person will tamper with few blocks, new blocks will get added and people will be he will never be able to reach up to the last block changing the hash value for the last block. So, that way we make it tamper proof the detail mechanism I will discuss later on as we progress in the course. So, this is just a kind of broad overview.

(Refer Slide Time: 13:48)



So, the second field which is there in the block header is the mining statistics.

So, the mine is to mechanism the generate hash. So, in case of bitcoin the hash function look something like this. So, you have the previous hash along with the set of transactions. So, this is the previous hash the hash of block k minus 1; the set of transactions and the random nonce value. So, the task of the miner is to find out this nonce value such that they can ensure certain difficulty on this generated hash value.

So, for example, the complexity in bitcoin is some something like that that whenever you are generating this hash value; you have to find out this nonce such that whatever be the value of H k it will have some 20 number of 0s in its prefix or first 20 bits will be 0s. So, that is the level of complexity

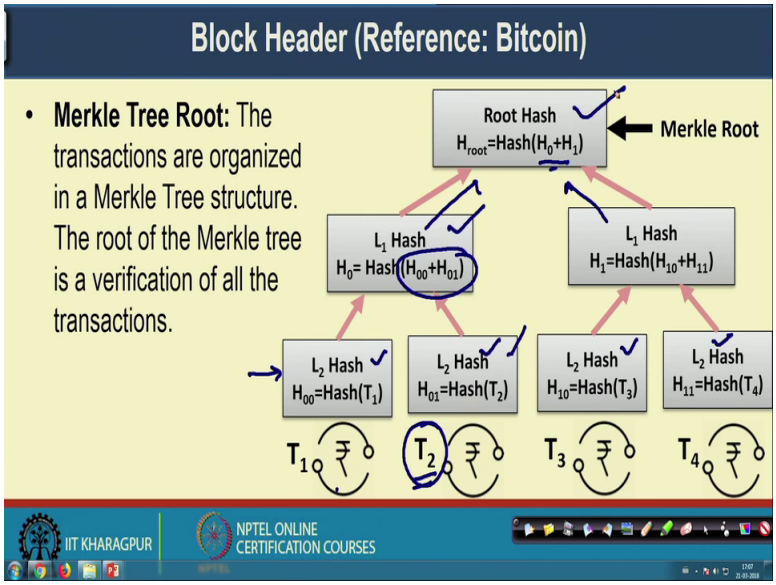
Now, by the property of the hash function if H k is known you will not be able to find out this message, but if this message is known then only will be able to find out H k. That means, what the miner have to do? They have to change this nonce, they will have to try

with different values of nonce to find out that when that objective is met the objective is that the generated hash value will have some certain number of 0s at the beginning.

So, that particular thing that how many number of 0s you want at the prefix or at the beginning that is called the complexity of the mining algorithm. So, the complexity of the mining algorithm will tell you that you want this many number of 0s at the beginning and as you increase the complexity of the algorithm you will require more time to find out this nonce value.

Now, this header blockchain header it contains this parameters, it contains the timestamp when that mining has been done the nonce value which is providing the corresponding hash value and a difficulty or the com complexity of the algorithm. The difficulty actually determines that how difficult it was to find out that particular nonce to meet the criteria of the complexity of having certain number of 0s at the prefix of your hash value.

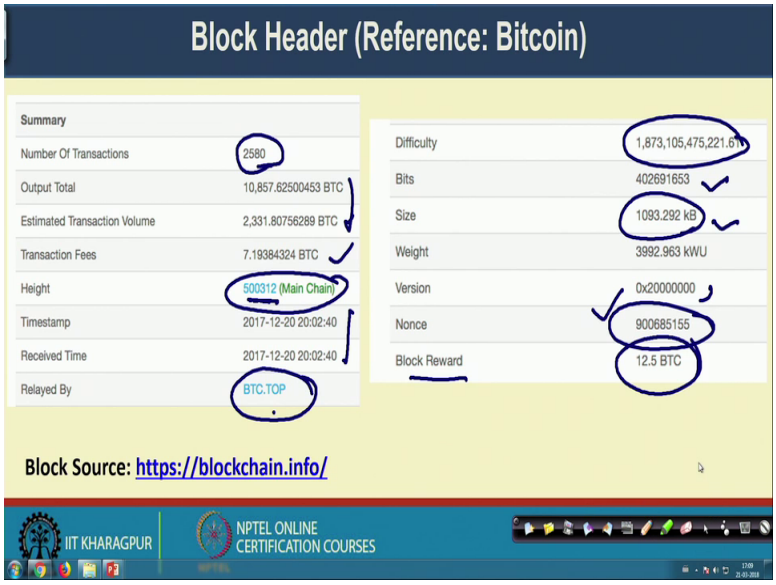
(Refer Slide Time: 16:07)



Now, the next part of block header it contains another parameters called Merkle tree root; what is this Merkle tree root? So, all the transactions in a in a block we arrange them in the form of a Merkle tree. So, I have discussed about the Merkle tree in the last class. So, it is like that at the root of the Merkle tree you have this hash of the transactions. So, every leaf node at the Merkle tree it contains the hash of the transactions and the intermediate nodes; they contains the hash of the combined hash values.

Now, the root contains again the combined hash values of its left tree and the right tree. Now interestingly if you want to make any change in the transaction say if you want to make a change in transaction T 2, then this hash will get change, this hash will get change and at the same time the root hash will get change. So; that means, if someone is changing one transaction; the root hash will change and once the root hash will change all the subsequent hash of all the blocks will get change because they are linked with each other. So, that is the beauty of the entire design of a blockchain which makes it as a tamper proof data structure.

(Refer Slide Time: 17:24)

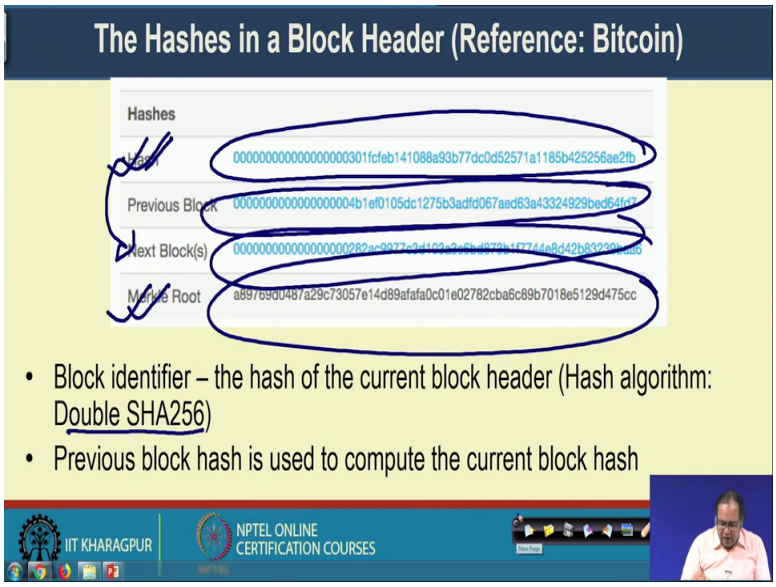


So, this is a typical block header the different parameters in the block header; here you can see that in that particular block the number of transactions which were there in the block the total amount of bitcoin that have been transferred the transaction fees that have been taken by the bitcoin network, the height of the block; that means, this is the height of the current main chain.

So, the chain that was there the block was at hash this number. So, the first block is number 0 the second block is number 1 that way this block is of number 500312 the timestamp values who has generated this blocks. So, this BTC dot TOP he is one of the minor who has actually validated the block. The difficulty level based on the hash finding algorithm, the mining algorithm that was there, how many bits was there in that block, the entire size of the block.

You can see that it is little more than 1 megabyte then some parameter called weight which actually balances among the size of different blocks; I will discuss that later on the bitcoin architecture version. The nonce which is used to generate the hash value; so, with this nonce the intended hash value has been found by this miner. And the block reward; that means, by generating this block what was the total amount of bitcoin that person has earned; now ok.

(Refer Slide Time: 18:56)

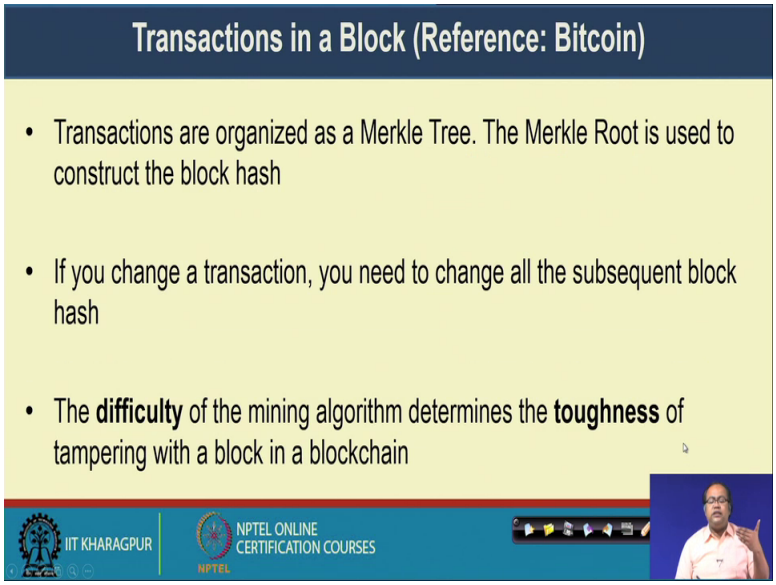


Now, as I have mentioned that the block header also contain certain number of hashes. So, these are the hashes; so, this is the hash value of the current block now this is the hash value of the previous block and if there are some next block; it also contains the shows the hash value of the next blocks although this is not contained in the header and this is the value of the Merkle root.

So, now if you are making any change in the one of the transaction the Merkle root get change; if the Merkle root get change then the corresponding hash value will get change, if this hash value will get change the next block hash value will get change. So, that way everything will get change and attacker have to actually change the entire blockchain. And initially I have told you that the current blockchain size is around 259; 250 GB. So, you have to make a change of that 250 GB of blockchain of size. So, the hash algorithm which is used to compute that hash it is a double SHA 256, I will discuss all this things

later whenever we will go to the details of this mining procedure and the consensus algorithm.

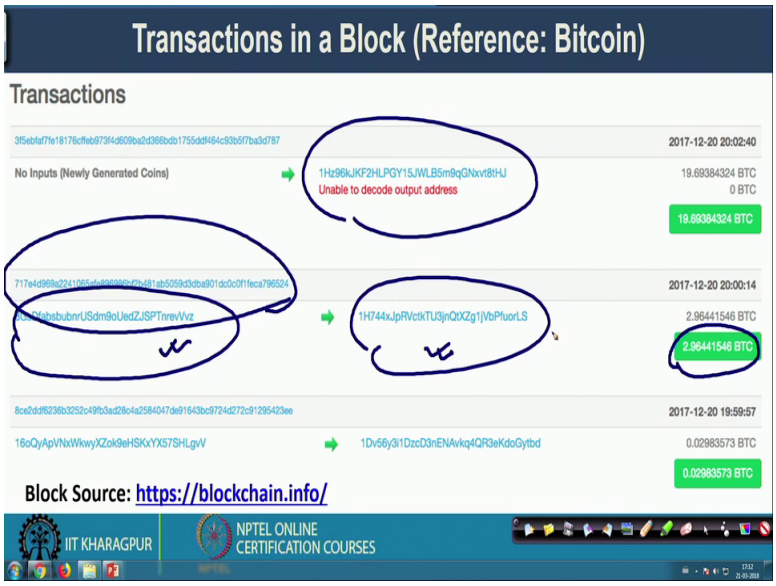
(Refer Slide Time: 20:11)



Now, the transactions in a block they are organized as a Merkle tree as I have mentioned and the Merkle root is used to construct the block hash. So, if you change a transaction you need to change all the subsequent block hash. So, the difficulty of the mining algorithm it determines the toughness of tampering with a block.

Now, you can see that if a attacker once to tamper with a block, he need to change the values of all the subsequent hashes. Now if the difficulties not very high then it may happen that by the time a miner will accepts a new block, a attacker can change the hash of all the blocks that way the attacker maybe successful. So, in this particular case you need to ensure that finding out that particular hash value is hash of hash the mine the attacker will not be able to change all the hash values in the complete blockchain.

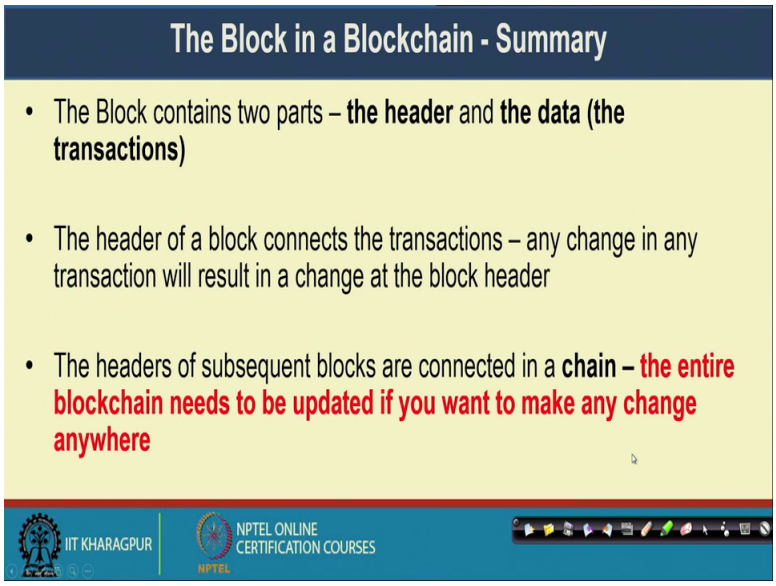
(Refer Slide Time: 21:12)



So, these are the transactions information inside the block. So, you can see that these are the users. So, here it is not a transactions rather this particular user got some amount of bitcoin by mining; so, that transaction is here. So, this is a valid transaction where this user has made a transaction to this user of around 2.96 bitcoin; so, that is the transaction in that block. So, what you can see here from here see; you cannot individually identify who is this user or who is this user, but you can determine that this is a kind of unique ID or unique address of a user who is making a transaction.

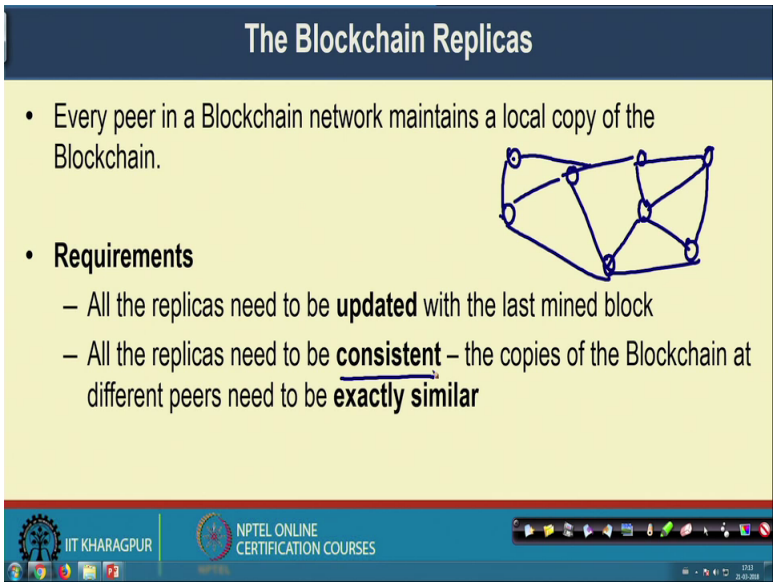
So, by looking into this transactions value you can validate the transactions and at the same time at the top you can see a hash value, this hash value is the hash value corresponds to this transaction. Now all these hash values are combined in a Merkle tree and finally, Merkle root is added as a part of the block header.

(Refer Slide Time: 22:16)



So, in summary that the block in a blockchain; it contains two part the header and the data the data contain the transaction the header of the block it connects the transaction. So, any change in any transaction will result in a change in the blockheader and the headers of the subsequent blocks, they are connected in a chain that way if you want to make any change in any of the block; you need to update the entire chain.

(Refer Slide Time: 22:44)

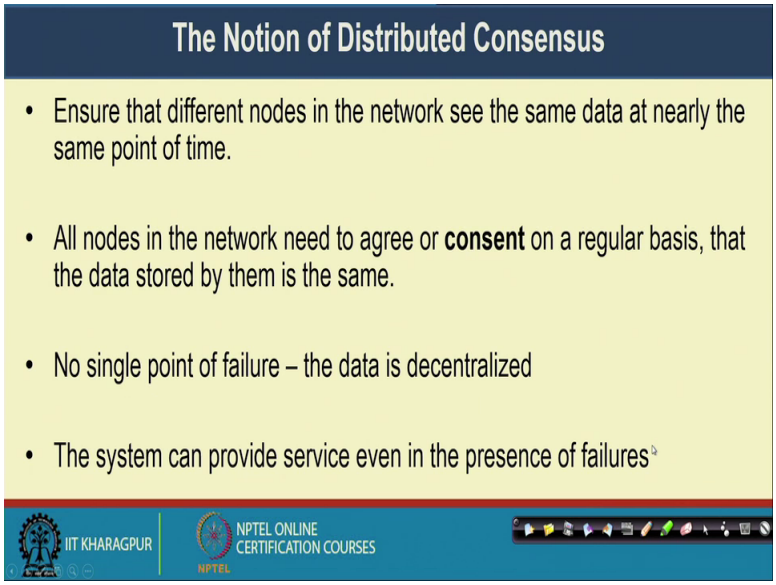


Now, the next thing is interesting part in blockchain is that how will you manage the replica? So, the idea of this blockchain is that there are multiple nodes in the network

who are interconnected and every node, they contained a replica of the blockchain. So, everyone of this node is maintaining a replica of the global blockchain.

So, the requirements are first all the replicas which are there at individual users they need to be updated with the last mined block. And all the replicas need to be consistent; that means, the copies of the blockchain at different peers need to be exactly similar.

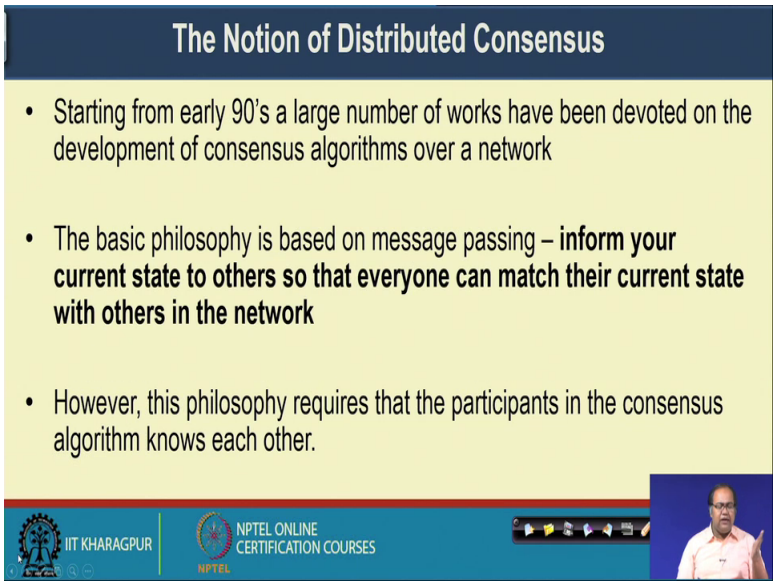
(Refer Slide Time: 23:28)



So, here the notion of consensus come into practice and a notion of distributed consensus where explode in the literature from early 1990's, where people have ensured that different nodes in the network, they see the same data at nearly same point of time. And in other words all the nodes in the network they need to agree or consent on a regular basis that the date of which is stored by them; they are similar they are exactly similar.

So, that particular algorithm we call it as a consensus algorithm and the consensus algorithm ensures that there is no single point of failure because your entire data is decentralized. So, if one node fails you have still have the data into multiple other nodes and so, the system can provide you service even in the presence of failures until and unless the network gets disconnected.

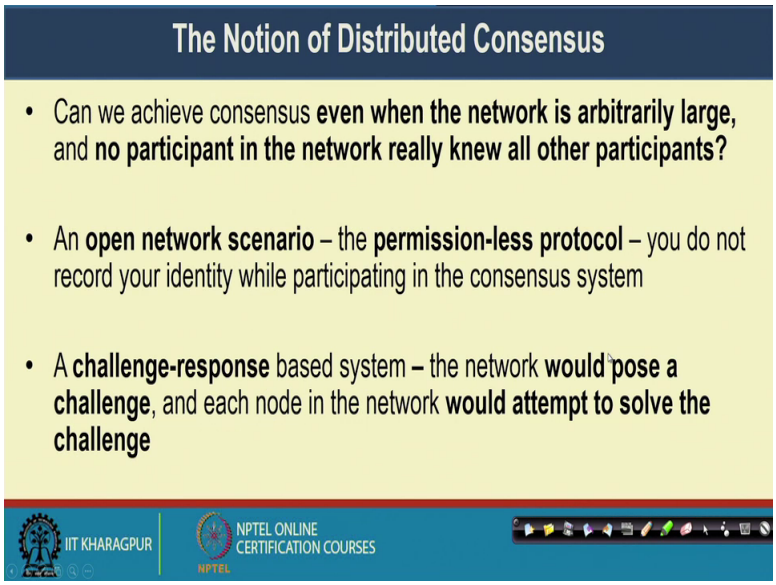
(Refer Slide Time: 24:22)



Now, as I have mentioned that starting from early 90's a large number of works they have been devoted on the development of consensus algorithms over a network. And a basic philosophy is based on message passing like you inform your current data to other nodes and everyone that way gets the data from all other nodes and the validate their local data. And that way you can see whether the data that you have whether it is a most recent data or whether that need data matches with the data of your peer.

Now, this philosophy requires that the participant in the consensus algorithm; they knows each other because you need to check or you need to find out that with which node you can validate your data.

(Refer Slide Time: 25:08)

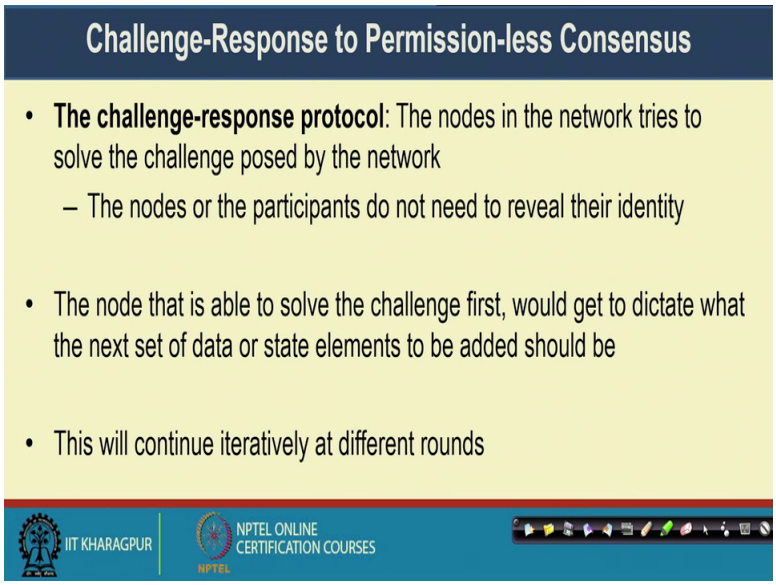


So, the questions question that comes in the blockchain network; if you remember the objective of the bitcoin network that bitcoin network was kind of permission less network. Permission less network in the sense that anyone can join the network anytime without reviling their authority; so, or reviling their identity.

So, in that sense the traditional distributed system algorithm based on message passing is not applicable here because you do not know that with which nodes you will validate your data. So, the question that we have here that can we achieve consensus even when the network is arbitrarily large; that means, no participant in the network really knew who are all other participants? So, we call it a kind of open network scenario or a permission less protocol.

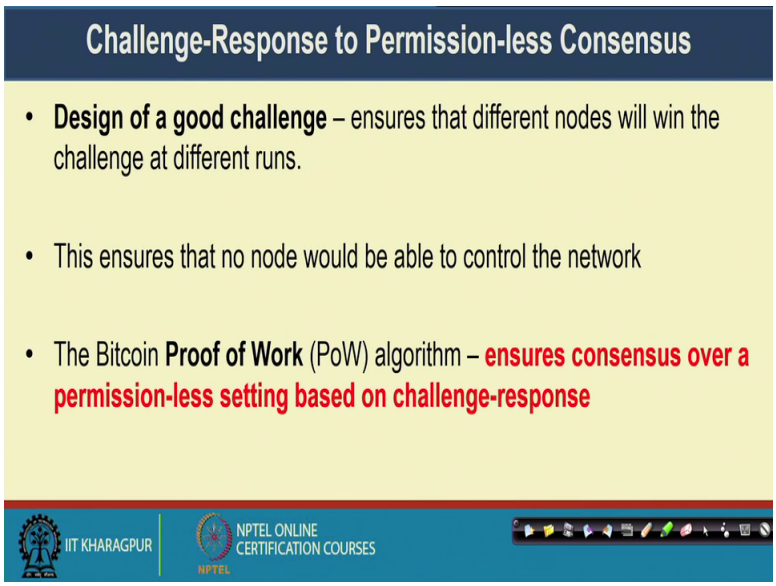
So, you do not record your identity while participating in the consensus algorithm, but still you will be able to reach in the consensus. So, to explore this kind of idea people have found out that well a kind of challenge response based system can work good in this architecture where a network would pose a challenge to the participants, every participant to will solve that challenge individually and each node will in the network would attempt to solve that particular challenge.

(Refer Slide Time: 26:23)



So, it comes to be a kind of challenge response protocol where the nodes in the network they tries to solve the challenge which is posed by the network. So, in that case the nodes do not need to reveal their identity. So, the network is giving them the challenge they have to solve the challenge and once they have to solved a challenge; they will announce that way I have able to solve the challenge first. So, I was able to validate the data you can add this data to the existing system; so, this will continue iteratively at different round.

(Refer Slide Time: 26:55)

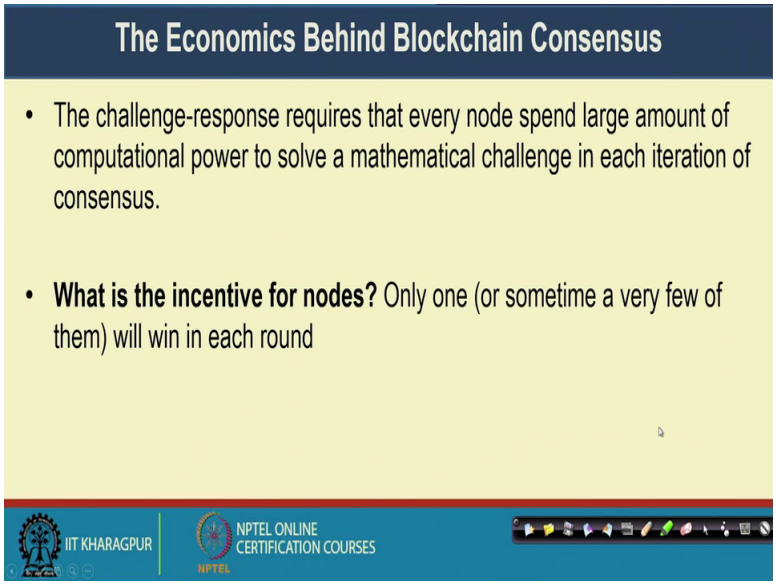


Now, the interesting fact is that if you can design a good challenge which will be posed at different rounds. So, you can ensure that that different rounds different nodes will win the challenge. So, that way your ensuring that no node will be able to control the network single handedly.

So, at one round one note will node will able to solve the challenge. So, that node will be able to say that way I am able to solve this particular challenge; so, this block is the valid block please add this block in the current blockchain. So, that was the idea which came into practice and this idea is known as proof of work algorithm in the context of bitcoin, which ensures that you are you are having a consensus over a permission less setting based on this kind of challenge response principle.

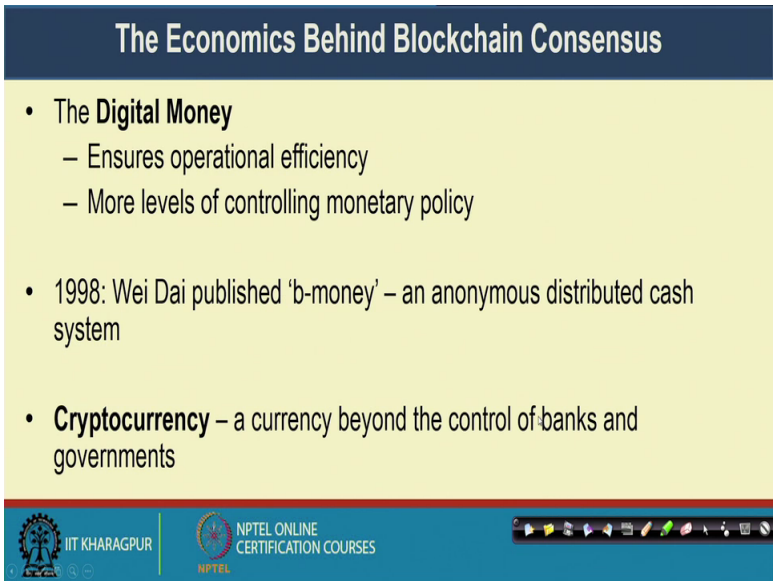
So, later on will discuss you details of the proof of work algorithm how it works, but this challenge response kind of thing is the basic principle behind proof of work algorithm.

(Refer Slide Time: 27:52)



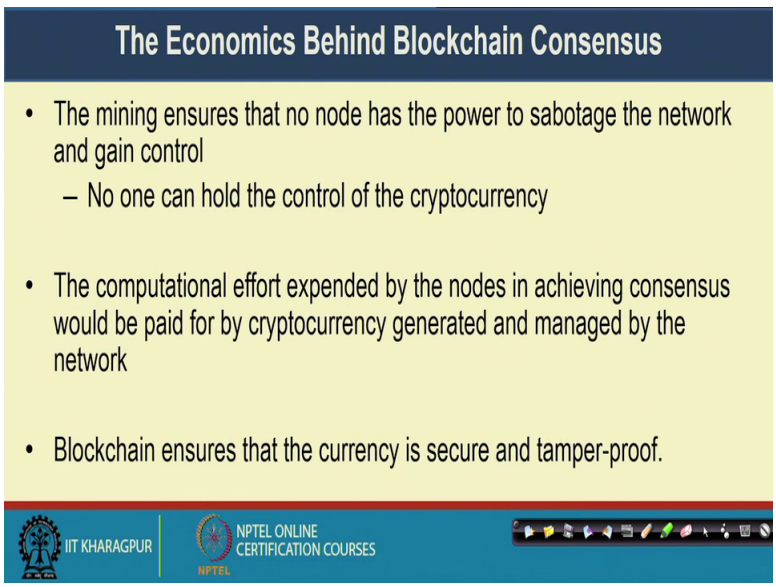
Now, there is another factor; so, the participate they are solving the challenge; that means, they are they are incorporating or they are spending a significant amount of resources like computation power, then time to solve that particular challenge and what is the benefit to them? What is the incentive for the nodes? Why they will participate in this challenge response algorithm? Because only one will win in each round, but others they are also spending their resources what would be the incentive for them.

(Refer Slide Time: 28:23)



So, here comes the concept of digital money which ensures the operational efficiency and it provides more level of controlling monetary policy. Now in 1998 Wei Dai; he published an concept call b money which is a kind of anonymous distributed cash system which is we can say it is the mother of this concept of bitcoin or cryptocurrency. So, cryptocurrencies something like it is a cryptographic currency there is no such physical currency. So, no one have to give the physical currency to a person rather than network will generate that currency; so, that currencies beyond the control of the banks and the governments.

(Refer Slide Time: 29:04)

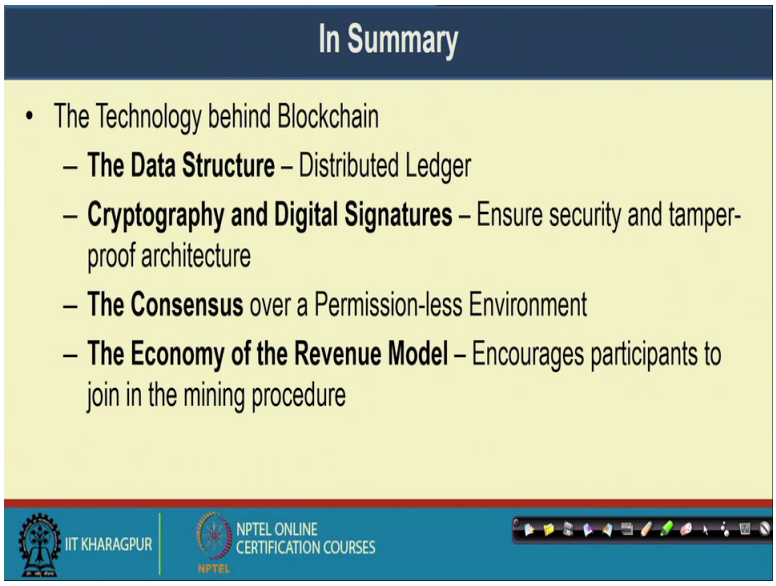


So, what you can do? That whenever the participants are there in the mining procedure, the mining ensures that no node has the power to sabotage the network and gain the control that is the good part of this cryptocurrency algorithm that no one will be able to hold a control of the entire cryptocurrency.

So, that was the basic philosophy like there should not be one centralized node like a bank or a government will control the entire monetary policy; at every round different people will able to add data to the blockchain. And the computational effort expended by the nodes in achieving the consensus algorithm, they will be paid with certain cryptocurrency which is generated by the network and managed by the network.

So, that way there is kind of monetary benefit to the miners that if they participate in the mining procedure and if they devote their computational resource and time in the mining procedure; they will get certain amount of money in return. So, that is the economic things behind this kind of bitcoin mining concept. So, the blockchain it is ensuring this currency is secure and tamper proof.

(Refer Slide Time: 30:12)



So, in summary there are multiple technologies behind blockchain and in this last 30 minutes I have given you a very quick overview of the entire thing. So, later on we will go to this individual things in more details; so, we have a data structure that minutes the distributed ledger which is forming the backbone of the blockchain.

We have the cryptography and the digital signature algorithms which is ensuring the security and tamper proof architecture of this entire blockchain data structure. We have the consensus algorithm over a permission less environment based on the challenge response scenario where you do not require to reveal your identity.

But still you can ensure that whatever data which is there at individual nodes they are correct data and then the economy or the revenue model behind this architecture that it encourages the participant to join in the mining procedure and to validate that a block is a correct block.

(Refer Slide Time: 31:08)

So, with this I will stop the this particular 30 minutes lecture. So, many of the part of the presentation I have taken from one of the nice block by Praveen. So, it was a block in Linkedin; Why so, Delirious about Blockchain; a technical overview. So, you are encouraged to read that particular block he has given a nice overview about how this concept of blockchain came into practice by exploring this different internal technical concept which is there behind the design.

So, in the next class which is the last class for this week; we will discuss about the different applications of blockchains and some internals about the bitcoins which makes it an interesting technology.

So, Thank you see you in the next class.