Introduction to Blockchain Technology and Applications

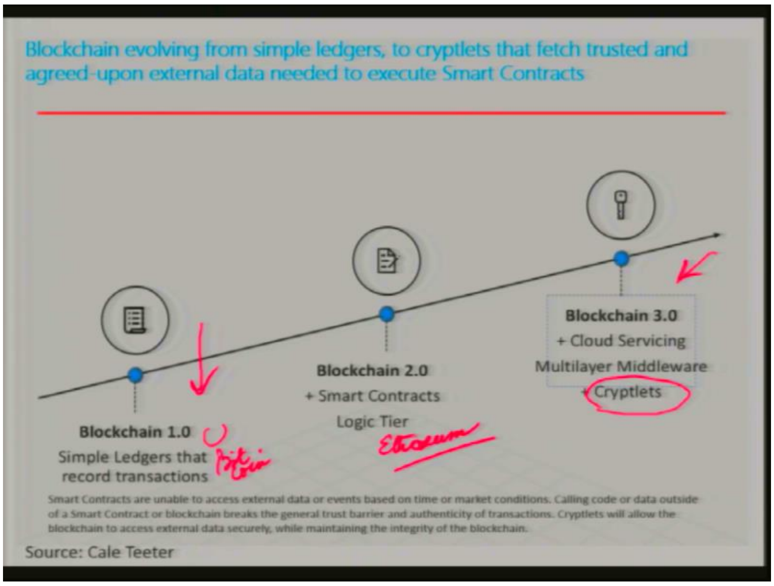
Prof. Sandeep Shukla
Department of Computer Science and Engineering
Indian Institute of Technology-Kanpur

Lecture No. 04
Blockchain Technology and Applications

Welcome to another session of blockchain technology and applications class. What we are going to do today is dive deep into the Bitcoin blockchain. In the beginning, I said that this course is not about crypto currency. But then why are we talking about Bitcoin blockchain? I have indicated before that Bitcoin is the first application of the blockchain technology, and therefore, it introduced a number of different technologies working together.

So it is more like integration of lots of existing concepts from cryptography from distributed computing and data structures. So therefore, it is no blockchain discussion can be complete without the discussion of how the Bitcoin blockchain works.

(Refer Slide Time: 01:10)



Also, if you look at this picture here, the blockchain technology is also evolving. So, what is happening today is that we have come a long way from the 2009 version of blockchain which was introduced with bitcoin. So, we normally call that as blockchain 1.0. So, blockchain 1.0 is basically simple, Ledger's that record transactions. And the idea of that recording of transactions was to make sure that any transaction that you make you ensure that it is replicated in many places, and also it is collected into blocks.

And then the blocks are made permanent by creating hashes of blocks and then storing the hash in the next block, and also making each subsequent block somehow dependent on the previous all the previous blocks by this hash chaining. So that is the blockchain 1.0. Then blockchain 2.0 was introduced with the idea of smart contracts. So, when we go and look into ethereum, later in this class, you will see that the blockchain 1.0 did not have a way of programming, how you want to use the blockchain.

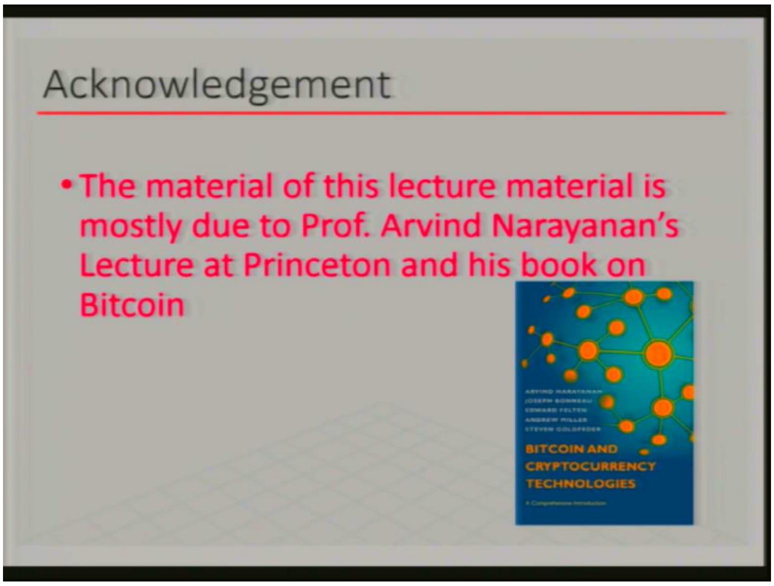
So, you can now in blockchain 2.0, you can actually write applications on top of blockchain that uses the basic primitives of the blockchain to accomplish data storage, transactions, and many other functions with smart contracts. So, that makes the blockchain 2.0 an example of that early example of this was ethereum and early example, of blockchain 1.0 was Bitcoin now we are moving into blockchain 3.0 so blockchain 3.0 is actually with the blockchain with smart contracts.

But it also has cloud, multi layer middleware, and also something called crypto wallets. Now, we will go into that much later in the course, if we have time, but the idea is that neither blockchain 1.2 or 2.0 can interact with the real time data that is coming from the from the outside. So, there are some tricks that are used in blockchain 2.0 to actually bring in real time data for example, stock prices that changes in real time.

Now in within, you know, 6 to 10 microseconds and that data every time you want to validate something, that data has changed. So therefore, there is a problem connecting real world data to the blockchain, especially for function, like a smart contract depends on what is the current stock price, for example, that kind of things might actually be a little problematic, and therefore, triplets are somehow designed to have a secure way of interacting with external data.

So, that is the evolution of blockchain technology. And that is where we kind of stand today. So therefore, what we want to do now is focus on blockchain 1.24 now, and understand how it works, and then we will move on to blockchain 2.0 and then finally, we will see some of the features of blockchain 3.0.

(Refer Slide Time: 04:47)



So today's lecture is actually based on Professor Arvind Narayanan from Princeton, his book and his lectures in at Princeton on a course on Bitcoin. So we will selectively use some of the ideas and some of the material that he used for his course. So we should acknowledge him.

(Refer Slide Time: 05:07)

So the first thing that we have to understand is the question of centralization versus decentralization. And we discussed this before that you have a way of collecting information like transactions, let us say in your bank account, and keep them in a one database. Or you can also have one replica of the database for fault tolerance or for backup, but the point is that the entire data is stored in a central location or in within the author authority of a central, trusted body, like your bank.

Now, that obviously has its own problem first of all, centralizing everything gives you this notion of, you know, vulnerability in the sense that if that particular data is somehow lost by a cyber attack or some other way, then you are going to lose all the information. Second is that the central authority has full command over what can be done on the data. And you have to somehow trust that central authority not to manipulate the data or delete some entries or do something like that.

So, for example, Aadhar is a perfect example of a centralized database. So all your biometric data, and everything is in central data's data centers. Some of them are in Noida, some of them are in Hyderabad. But the point is that if you have your other number and biometric, then people actually use API's to access that data to authenticate on authenticate you that is through the KYC process, when you use your other based KYC.

Now, if the central authority maintaining the database decides to delete your entry, then you no longer have any access, nobody will be able to access your data and therefore, you will be not recognized through the KYC process. Now, unless there is a alternative way of doing KYC, you will be in trouble in the sense of a sense that it is not it is denied that you exist. So, so that is a problem of the centralization, too much power in the hand in the hands of a central authority.

So, therefore, now, Satoshi Nakamoto he was very worried about or he or she or them, they were very worried about these centralization of monetary infrastructure and monetary information in the hands of few banks and 2008 there was a big banking crisis in the US and, a lot of banks actually went bankrupt and government had to bail them out because they were too big to fail.

And therefore, he was saying that we have to deliver ourselves from the tyranny of the banks, and then we should actually create a currency that is completely decentralized. So, there is no central authority which creates that money or a central authority, which, keeps the keeps track of the money or authenticate people to use their money. So that is the basic, so, you know, political economic underpinning of the introduction of the Bitcoin blockchain.

So centralization and decentralization versus decentralization is a basic concept that underlies the blockchain technology. Now, centralization has many advantages.

(Refer Slide Time: 09:13)

For example, it is easy to manage. So a bank will not have to worry about, you know, having replicas and privacy issues. Because if you have many replicas, you have to make sure that the data is kept private, all that kind of stuff. It is easy to provision. So for example, if you want to create a new account, you have to just make change in the central database. If you want to delete an account or then also you have to change in one place, it is easy to ban.

So for example, if you have a centralized authority, maintaining for example, your DNS domain name service, then you can easily ban certain domains from being found, and it is easy to distribute responsibility. So centralization has all these advantages. However, decentralization has some of these disadvantages, for example, it is harder to manage if data is distributed all over the place.

It is harder to distribute work in the sense that you cannot say, you know, you have this responsibility, and you have that responsibility, because nobody is being commanded by a central server or central authority. And it is harder to ban something so you cannot delete information so easily. And it is harder to provision. But centralization, as a single trusted party, is its biggest weakness in terms of not only the fact that they might actually do things in a unilateral way, which is not good for the consumers.

But also the fact that if it gets cyber attack or any kind of attack, then also it can actually have a devastating consequences. Whereas if the data is decentralized, and all the functions are decentralised then it is much harder for an attack or for a central authority to you know take over dysfunctionalities or change the functionalities and so on. So, therefore, blockchain is based on the notion of trust.

So, the question is when you have a central authority, you have to trust that authority and normally in the past we have been doing that. So, we have been trusting the authority that is managing the DNS, we are trusting the authority that manages the digital certificates, we are trusting the authority that is managing our accounts. But in the current time, we have been, we have we are seeing that a lot of the times the central authority becomes overly authoritative, they ban things.

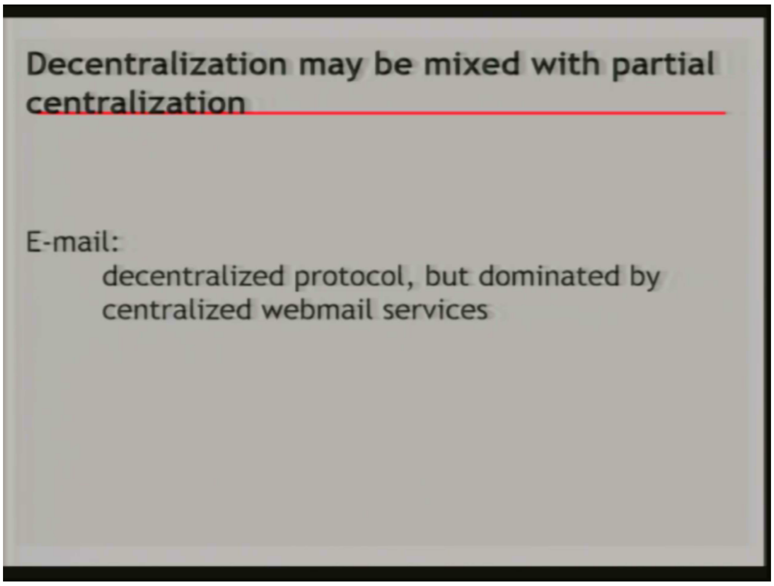
They actually can change things that need to put you into trouble too therefore, the trust has to be created from a decentralized system. Now, in a decentralized system when you have many players and no player is more important than the other players, which is the case in the centralized situation, there, creating trust is also a challenging task. Because you have you may have a number of the players in the system a number of actors in the system, who are also malicious.

So, you have to somehow assume that in your system in the ecosystem who are actually together doing the computation together, keeping track of data together deciding what is valid and what is not valid. You have to assume that some of them will actually do something maliciously and so, therefore, your system should be designed in such a way so that in spite of that you have the ability to trust the system.

So you have deriving trust from untrusted participants or an untrusted actors for that you may have to make certain assumptions. For example, what percentage of people are untrustworthy? And what percentage of people are trustworthy if that assumption sometimes have to be met in order to trust the entire system. If you did not trust anybody, 100% of the people, then you cannot build a trusted system out of 100% untrusted actors.

So we will see how all these assumptions play out in the blockchain, especially in today's class, we will talk about bitcoins blockchains trust how do we how we derive trust in the system. Now sometimes, you know, it is not a black and white thing.

(Refer Slide Time: 14:02)



So we may have decentralization, but we may also have partial decentralization or partial centralization. For example, email system, the SMTP protocol is a completely decentralized protocol. So we can have SMTP servers. And the email client that we use are SMTP clients. So you have all these clients. And we have, we can have many, many SMTP servers who talk to each other. So if my simply when I send an email, my SMTP server then talks to the other SMTP server in the destination.

And then that SMTP server will actually forward the mail to the SMTP client that is the let us say, outlook or whatever, to read that email. So this is a completely decentralized protocol. It was designed like that however, in the past 20 years or so, almost 30 years now, since the introduction have Hotmail, Yahoo Mail, and then now Google Mail, and Microsoft mail and so on, we see that certain all mails eventually go to the SMTP server of those 4 or 5, 6 different web mail servers.

And then from there, everybody gets the, everybody is received that email. So I your, if you are using Gmail, for example, and your sender is also using Gmail, then they both are talking to a single SMTP server, and not like their local SMTP server talking to the local SMTP server of the receiver. So that is a decentralized protocol that is designed, but for convenience, and for better management people have found kind of a centralized way of managing the entire email infrastructure.

So this can happen on their centralized infrastructure or, for example, when you buy a digital certificate, you buy it from a Certification Authority. So there is a usually a Certification

Authority, which is, which may be, affiliated to a larger Certification Authority, but in general that Certification Authority issues and signs that digital certificates so when you buy the digital certificate you are from a digital certificate issuing authority, then what you get is a digital certificate signed by that authority.

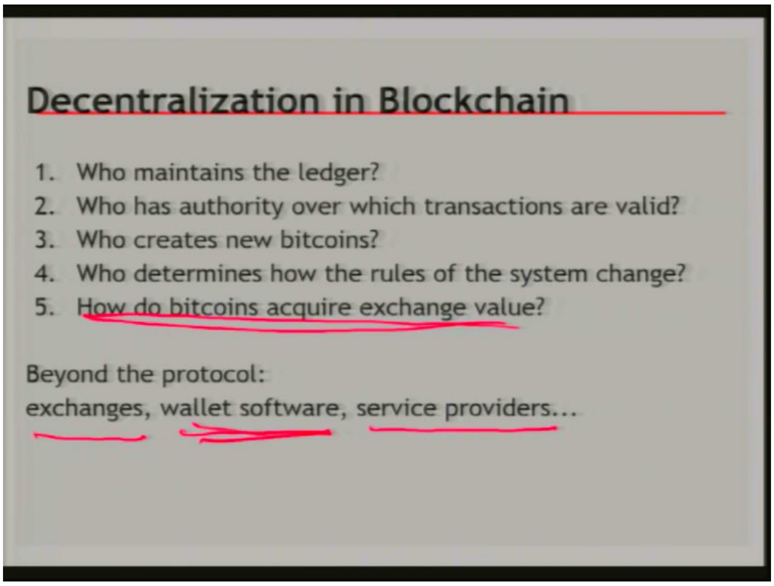
And then that authority has to be trusted by the person who is actually going to use your digital certificate to, for example, send you encrypted message or check your signature, things like that. Now, if that or authority, then becomes annoyed with you and then decide to revoke your certificate, then you can no longer use your digital certificate because the other entities will not be trusting that that digital certificate that you have because it has been revoked.

So here again, the central authority will have too much power to actually control whether you are trusted by others, or you are able to digitally sign documents, things like that. The other central example centralized example is DNS. So DNS is the domain name server. So domain name server like when you have a URL, for example, let us say www.google.com. Your browser then asks the DNS server who is, what is the IP address, for that for that domain name because the internet protocols did not understand domain names.

They only understand numeric IP addresses. So it goes to the DNS server. It has a hierarchy which is there is local DNS server there is original DNS server, there is a domain, top level domain, DNS server and so on. But eventually it goes all the way to a central root server. So in this case also suppose somebody decides, the companies that manage this DNS servers decides that google.com is no longer a desirable domain.

If the authorities that maintain the DNS servers decide that google.com is no longer a trusted, trustworthy domain name, then they can revoke it, and therefore, the Google will have a problem. So, these are the kind of things that can happen in a centralized server based system. So, that is why the decentralization comes into picture.

(Refer Slide Time: 18:52)



Now, how is blockchain decentralized? So let us ask a few questions. And if you can answer these questions, then you will know whether it is centralized or decentralized, for example, who will maintain the ledger. The ledger is basically the transaction records. Now, as we saw in the previous exercise that we talked about before, that we created a ledger, which has all the ingredients of a blockchain except that the ledger was maintained by the program that you were writing. So there was a centralized ledger, even though other components of the blockchain was there.

However, if you do that, then that centralized ledger, centralized program, which is maintaining that ledger will have all the power to change things and stuff like that. So therefore, in blockchain, we did not maintain the ledger in a central only one location. It is replicated among 1000 and 1000 of different computers and different what we call nodes. Now who has authority over which transactions are valid?

So, in a regular banking situation, you see that the transaction validity is checked by the bank in the program that we have done as an exercise. There also my program or the program you will write was taking the validity of the program what validity of transactions. Now, we said that the validity of the transaction has this property that amount in the balance has to be greater than the transaction amount greater than or equal to the transaction amount.

We said that the detail signature must check out all this stuff but if it is only one authority that is checking it, they might surreptitiously change the logic they might say, I will allow now transactions which has which is which will end up in negative balance from the sender's account or I will not allow this person signature because I did not like this person. So, centralized authority for transaction validity checking is also not good.

So, in blockchain, actually, every participant who is interested in data mining or keeping the ledger is going to check the transaction validity. So, if somebody wants to act maliciously, but others will not necessarily act maliciously, at least, the majority will not act maliciously we assume. And therefore, the transaction validity should be within the logic that we have come up with. Now who creates new bitcoins?

So we know in a banking system, the new currencies created by the central authority like Reserve Bank of India, or Federal Reserve in the case of in the United States, but in the bitcoin blockchain, the bitcoins are not created by any central authority. In fact, we discussed before that whoever actually wins the competition to create the new block. And if that block is actually accepted as the new block, then they get some reward.

And that is the only way to create a new bitcoins. So therefore, it could be anybody who wins and not every, time same person wins. Therefore, bitcoins are created by many, participants. Now, who determines the rules of the system change? So bitcoin has now certain system rules that everybody is supposed to follow every participant. So that is how they program the program they write to run the mining process or check for transaction validity is based on certain rules.

Now, how do you change the rule? The question is that you might want some rule change because you found some deficiency in the system. And that is when the rule changes. When rule changes the current bitcoin current part of the blockchain that so far you have developed may become obsolete. So, you may have to fork the blockchain. So, that is also not done centrally because this forking process takes place kind of automatically through the entire dynamics of the system.

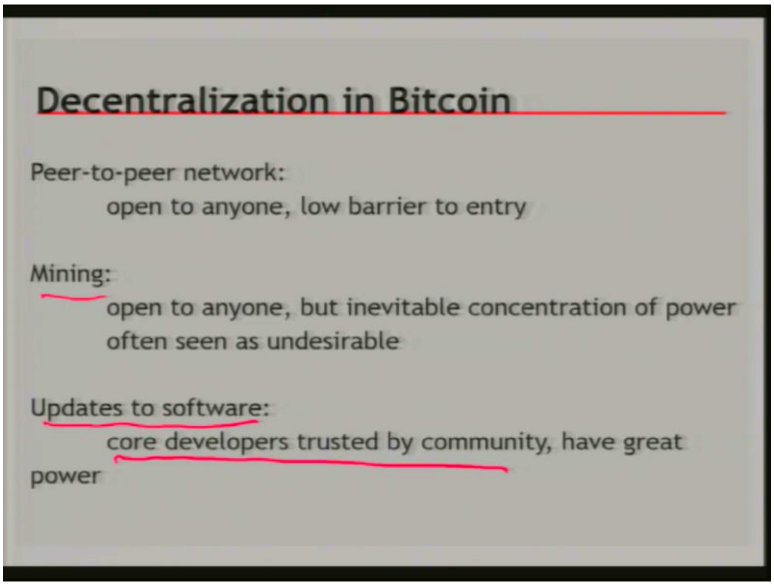
Now, other question is how do the bitcoins acquire exchange value that is, what is the monetary value the fiat currency value of bitcoin? So, last time I checked, I think, this week, the bitcoin prices about is between 7000 and $8,000 per bitcoin, there was a time when it was only a few 100 dollars, and there was a time when it almost touched 20,000. Now, who decides this exchange value there is no central authority here.

But this exchange value is not part of the blockchain dynamics or the blockchain programming or anything it is done external to the bitcoin blockchain it is done by the market, but the market works also in a decentralized fashion because more people if more people get interested in the bitcoin buying bitcoin, then its price goes up, if people start feeling that bitcoin is may have a problem.

Then for example, when news comes that China is going to ban bitcoin or things like that, then the picture bitcoin prices go down. So this entire going up and down is based on how many people are trying to buy bitcoin or how many people are trying to sell bitcoin. If more people want to sell bitcoin, then the price will go down. So that is how this process works. So these are actually this thing is actually beyond the blockchain protocol.

This is actually completely extra external to the blockchain protocol. So the bitcoin exchanges, the wallet software, some wallets software is the software that you if you are participating in the bitcoin in buying and or doing transaction with bitcoin, then you have a wallet software, you have various service providers. So these are external to the protocol. So we will not worry about these in this course.

(Refer Slide Time: 25:19)



So how is decentralization achieved in bitcoin? First of all, bitcoin actually works with a peer to peer protocol. So if you make a transaction, you actually want to tell the entire ecosystem that I am making this transaction. And here is my signature authorizing this transaction. And this is broadcast, how do you broadcast it, you broadcast it by sending it to all your local neighbors, and then all your local neighbors, then will send it forward to their neighbors.

And this is how the blockchain transactions are broadcast throughout the entire network. And this entire network is actually overlay network over the TCP IP internet and so on. So, therefore, in this network, the nodes or whoever is participating in the bitcoin ecosystem. And the communication is done through an overlay over the TCP IP because the bitcoin endpoints are basically applications, programs that basically do all the work.

So, anybody who wants to say that I want to join the bitcoin network, he can just have the software that is required to join the this peer to peer network and he can enter and the way he enters he has to tell what is this public key, actually the hash of the public key as we discussed before, as an address, so, he says that here is my address, and here I am. And here is my public key. So then the rest of the world eventually would know that there is another search node that is from this part of the network.

Now, the mining process that is mining is a process of deciding, which is the next block that will be added to the blockchain. Even though the 10s of 100s of copies of the blockchain at every node, there is 1 copy, but the copy is dynamic so, because it keeps growing. So, what gets added next has to be agreed on by all these different entities because otherwise the blockchain will be looking different at the different nodes.

So, how this process works is to what you call mining, and anybody who wants to do mining can do it so, if you want to do mining, you can also do it. However, it turns out that today, doing a bitcoin mining requires you to solve a very difficult hash puzzle. And we saw before that solving a hash puzzle requires brute force, compute and therefore, you need a huge amount of computational power and only a few players seem to have acquired that much computational power.

So, with a single laptop or a single workstation, you cannot really ever win the mind the hash puzzle contest. So, therefore, you will never be actually mining any block. But if you have millions of dollars worth of equipments and software, then you can sometimes win the blockchain this hash puzzle and therefore, you will be mining and you will get rewards, the question is, who gets rewards?

See, third thing is that the updates to the bitcoin software, sometimes you may decide that the rules have to change and we are there you find a bug in the software, which everybody else is running. So, therefore, there are some core developers, which are who are trusted by the community and they have to announce a new update, everybody has to get that update and if the update is very, you know, changes the way validation has to be done or mining has to be done.

Then there will be a big change in the blockchain. And maybe sometimes blockchain may fork. So we will stop here. And then when we come back, we will we will actually go into the idea of consensus mechanism in bitcoin blockchain.