EPISODE 537

[INTRODUCTION]

[0:00:00.3] JM: Decentralized applications might someday offer alternatives to modern monopolies. Uber, Airbnb, Facebook, Amazon, all of these services could be re-created on a decentralized stack of technologies like Ethereum, IPFS and Golem. Fully decentralized services could be more transparent, cheaper and more efficient, but let's be realistic. Today, even the simplest applications of fully decentralized block chains don't work as well as we need them to. CryptoKitties offered a glimpse into how a simple viral application can limit the throughput of Ethereum, and don't forget that these technologies are someways still subject to centralization in their current form.

Miners form the decentralized consensus layer and that mining activity is physically centralized in large mining farms. The decentralized future is possible, but in order to get there we will need to make progress on the low level tools that such a world will eventually be built upon, and this is the realization that today's guest, Carl Floersch, had. Carl is a researcher for the Ethereum Foundation, and he was initially excited about the prospect of decentralized apps, such as a decentralized Uber. But as he looked more closely at the space, Carl realized how early we are and how much work there is to be done on foundational technologies.

Proof of stake is the central topic of today's discussion. Proof of stake is a consensus mechanism that is an alternative to proof of work. In proof of work, miners race to validate blocks of transactions. This results in duplicated effort, and perhaps wasted energy. In proof of stake, validators are chosen to approve transactions. These validators lockup and amount of currency that they are willing to stake. If a validator acts badly, the validator will lose their entire stake. This mechanism could be more efficient than proof of work, and we will explain why that is in this episode.

If proof of stake does function, it could lead to a faster, truly decentralized Ethereum blockchain, and that's a remarkable potential outcome. If you're interested in checking out our open-source community, you can go to github.com/softwareengineeringdaily or you can download our apps to check out what we're working on. You can check out softwaredaily.com to find those apps.

We've also got meet ups coming up in New York and Boston and also in LA eventually. So if you want to follow our meet ups, you can go to softwareengineeringdaily.com/meetup. With that, let's get to this episode.

[SPONSOR MESSAGE]

[0:03:03.6] JM: Today's sponsor is Datadog, a cloud scale, monitoring and analytics platform. Datadog integrates with more than 200 technologies so you can gain deep visibility into every layer of your stack and any other data that you're interested in tracking as well. For example, you can use Datadog's Restful API to collect custom metrics from your favorite crypto data sources and analyze trends in Ethereum prices over time.

Start a 14-day free trial, and as a bonus, Datadog will send you a free t-shirt. You can go to softwareengineeringdaily.com/datadog to get that free t-shirt, and thank you to Datadog for being a continued sponsor. Get that free t-shirt at softwareengineeringdaily.com/datadog.

[INTERVIEW]

[0:04:01.1] JM: Carl Floersch, you are an Ethereum developer. Welcome back to Software Engineering Daily.

[0:04:05.6] CF: Hello, hello. How is it going?

[0:04:07.8] JM: It's going great. It's really good to talk to you. Last time we spoke, I was still very unfamiliar with decentralized technology, and I was interviewing all these people where at the end of the interview I would be thinking that this person that I just interviewed is very intelligent, but this technology is totally alien to me. I didn't know why these crypto people, such as yourself, wouldn't just go back to building web apps and leveraging Amazon, but now I get it, actually, because after spending a little bit of time with this, I get that there are actual technological breakthroughs. What was it that made you commit to the decentralization space before there was all these hype, before there were so many people involved? What was it that clued you in that there was something important going on here?

[0:05:03.2] CF: I got into the space essentially because I have for — Ever since I grew up, my dad has been a bit antiauthoritarian and alternative and I resonated with this kind of decentralized future. I didn't really know what it was. I knew that I was searching for a kind of re-envisioning of how we might be able to even structure society in a broader sense. Is it capitalism? Is it socialism? Is it communism? Is it libertarian? Whatever. I saw this technology and I definitely didn't know what it was, but I felt like there was this kernel of something different, something that is fundamentally new about it, and it's taken me. So I kind of went on that, went on that gut feeling, and it has brought me down this crazy, crazy rabbit hole and I think I'm starting to get an idea of what that is and like what decentralization really means.

So now when you ask me this question; why am I in the space? I can say that it's because I think that there is a lot of value in constraining central authorities, constraining power. So what that basically means is if you ensure it, absolute power corrupts absolutely, something along these lines. When we have this ability to constrain power, it actually gives people a level of trust that is not able to be achieved in other settings, because you know that there is some recourse for bad behavior. So this has been my focus from then till now. Now I just kind of have a better idea of what it really is that makes this space special.

[0:06:49.6] JM: Yeah, because I remember the last time I talked to you, most of what we discussed was the big picture, and I think this was largely because I was not super familiar with the engineering side of things in the decentralization space, but you are very good at explaining the big picture; the decentralized Uber, the decentralized Airbnb . Was there a moment where you realized that there were some actual engineering underpinnings and that we weren't just telling ourselves a story about how we would eventually get to this decentralized future?

[0:07:20.2] CF: Yeah. I would say it was an exploration into what a like smart contract blockchain can do. I started out, I knew that you could create a decentralized Uber and you could do it by — It seems clear what those rules were. However, I didn't know what the actual design principles and architecture philosophies came in or behind this technology. So the thing that really allowed me to realize this is working at the Ethereum Foundation, working with Vitalic on Casper and Vlad. Casper, by the way, is Ethereum's proof of stake protocol replacing proof of work. We could get into that. That's a whole long story. But working on that, I realized that there is a fundamentally different way to design these applications and that means that we take

into consideration economic incentives. It's this concept of crypto economics where you're able to say, “Okay. What is the essence of decentralized Uber, because we’re using that?” It's essentially taking Uber, but we are replacing parts of it that central authority that currently are able to, for instance, like price gouge the drivers or engage in unfair, tipping or something like this, and that is actually something — That is a design practice in the domain of software engineering. I'm just currently — I think everyone is currently learning about and getting a better understanding of.

[0:08:54.9] JM: Yeah, and I have a bunch of questions about Casper. I did my homework a little bit better this time, but just a few more questions about this area though to warm us up. Since you mentioned, your dad had an influence on you and you've been thinking about this stuff for a while. It does feel like the pernicious aspects of centralization are coming home to roost recently with kind people starting to feel anxious about Google, people feeling anxious about Facebook and Twitter all for various reasons, but somewhat related reasons, sort of this opaque censorship or potential abuse of monopolistic power.

I've read this article recently about the case against Google. I don’t know if you saw that one, but it was just about how Google penalizes vertical search engines and prevents any search company from coming up allegedly. One thing I wonder about is if you have a sense that there needs to be any kind of regulation — Because what I love about the decentralization world is you can look at this and you can play this out into the future in like 5 or 10 years and it doesn't really seem like there's anything that a Google, for example, could do to obliterate the world of decentralization coming up with basically a full stack that is an alternative to the Google world.

Then you've got no hardware manufacturers that could come up with a hardware solution that would have properties that would take advantage of that decentralization. I think I remember seeing — Actually, I heard a comment from Vitalic, the Ethereum founder, and he said his main concern — He's pretty optimistic, but his main concern is Google working with the US government to come up with like a cryptocurrency that everybody just defaults to, which I guess that would be problematic. That would be potentially like, “Oh! That's worrisome for the decentralization space,” but I just think it's funny because it seems like the Ethereum folks are the ones who have the best chance right now of building an alternative to the Google world and

they don't seem to even think about regulation. They don't seem to care about regulation. They're like, "Let's just build something better."

Do you feel like regulation is something worth considering in our centralized world?

[0:11:18.2] CF: I think that regulation absolutely will shape the ecosystem moving forward. So a restrictive regulation may really hinder a lot of progress and it may hurt the development of this space. So I definitely do think that regulation plays a key role.

[0:11:38.1] JM: You're thinking about like regulation of ICO's or stuff, like regulation against the cryptocurrency space. That's what you're thinking about.

[0:11:44.9] CF: That's what I was thinking about. What regulation were you thinking about? Because there were multiple things there that I definitely would like to touch on.

[0:11:49.9] JM: Of course, yes. Well, so I guess there's two — Certainly, two questions. Well, I was thinking more in the everyday conversation of, "Do we need to regulate Google?" Do we need to regulate Facebook?" Because they're abusive monopolies that are preventing people from starting rival businesses. I guess that's one side of things. I hadn't even thought about bringing up the ICO or cryptocurrency regulation space, but you can feel free to address those separately or in a way that ties together, if you like.

[0:12:16.9] CF: Okay. Yeah. So there is — Regulation is in some ways the same or similar domain as cryptocurrency, and I'm talking about not regulation of like Google and Facebook, etc. Regulation is a constraint on some kind of authority, some kind of power. Where you are describing Google has these monopolistic business practices allegedly, and a response to that could be, "Okay. The government needs to, now, regulate Google."

However, I believe, especially these days with technology moving so rapidly, a better option is very likely going to be cryptocurrencies and the kind of self-imposed regulation in these design principles, the constraint on power. So if I were to design a Google, maybe there would be some economic incentive which prevents this monopolistic power.

Now, at the same time, you are absolutely right. You touched on the Vitalic's fear of the United States government and Google coming together to create a currency. I would say that the biggest threat to cryptocurrencies right now most likely, I mean, the one that resonates most with me at the very least, is this idea where a decentralized technology, this decentralized tech, which has enabled a different kind of trust than we've ever seen. That kind of gets co-opted and corrupted and re-centralized, but packaged as a decentralized platform. There is the possibility of a future where we get most of the decentralization that, for instance, Ethereum is working to provide. But we get most of the trust that people are using from these platforms. In other words, the ability to deploy a smart contract and know that it's not going to be censored or whatever, but then there is still like the master keys held by some central organization. That would be super scary, because there's so many new people getting into the space. There is the possibility that a centralized system that claims to be decentralized is actually used, or may be that the — I don't really need to go down all the different hypotheticals, the thing that Vitalic said being one of them. But this is definitely scary, because the thing that I'm excited about is the new design principles, the new architecture, the new web stack that we can create and that is crafted such that central authorities do not retain absolute power. That's the exciting part. If we convince people that, "Oh! No. Essential authorities don't have absolute power, but in fact they do have some kind of bad sway." That gets really dangerous, because then you get the oppressive state that you don't even know is oppressing you.

[0:15:15.8] JM: The 1984 scenario.

[0:15:18.2] CF: Yeah.

[0:15:18.8] JM: I think this will have a relation to our conversation, because the way that Bitcoin, well, I guess, and Ethereum, if I'm correct, are arranged right now is you basically have an AWS type of entity that is managing the consensus, which is the collection of mining organizations where all of the validation work is being done, and if somebody took control of all those miners, you could really disrupt the network. Is that accurate?

[0:15:53.0] CF: Yeah, absolutely. I could continue. This is my specialty in a lot of ways.

[0:15:57.8] JM: Please. No, please. We'll get in to the mechanics of proof of stake and why that is, perhaps, an alternative that leads to a more pleasurable form of consensus with less dangerous centralization. Yeah, I mean, go ahead and address that point from a high level.

[0:16:16.1] CF: Yeah. I mean, from a high level, of course, the point of these systems is that, "Okay. You design some incentives, and these incentives in under normal circumstances or under even extenuating circumstances will produce results that you find desirable, and that means in this specific case, creating a single blockchain, a single record of history that is not censorable. In other words, you can't censor new transactions or that cannot be reverted. So you can't actually buy a house and then have your house — The money kind of reverted or the record reverted, either one.

So these are the properties that we want to maintain with proof of work as it is today, and you're absolutely right, that a system that is designed architected to be decentralized doesn't necessarily mean that it is un-corruptible, and that is actually kind of the point, this is kind of what I was getting at with saying that, "Okay. There's something that they can masquerade as a decentralized platform," but in fact to be a centralized platform. Yes, we are working with proof of work today, and as you said, we're working also to replace it with Casper.

[SPONSOR MESSAGE]

[0:17:31.5] JM: Software Engineering Daily is brought to you by Consensys. Do you think blockchain technology is only used for cryptocurrency? Think again. Consensys develops tools and infrastructure to enable a decentralized future built on Ethereum; the most advanced blockchain development platform.

Consensys has hundreds of Web3 developers that are building decentralized applications focusing on world changing ideas, like creating a system for self-sovereign identity, managing supply chains, developing a more efficient electricity provider, and much more. So, listeners, why continue to build the internet of today when you can build the internet of the future on the blockchain?

[INTERVIEW CONTINUED]

[0:18:47.8] JM: So proof of work is a mechanism we've explained several times in recent episodes, but to abbreviate it, you have all these transactions that people are making throughout the world and then the different miners have pools of transactions and they pull from those pools of transactions and they are racing to solve a cryptographic puzzle associated with a collection of transactions at any given time. When they solve the cryptographic puzzle, they have validated a block of transactions and they broadcast that to the rest of the full nodes on the network and it propagates to the light nodes, and that's how you come to consensus about a blockchain in a proof of work scheme.

How do things work — and this is basically how things work in Bitcoin and Ethereum today. How does a proof of take system contrast with that proof of work-based system?

[0:19:45.2] CF: Great questions. So proof of steak and proof of work both rely on this limited resource. There is something that is limited, and that the resources used as essentially voting rights for what is the main chain. So the limited resource and proof of work is energy. We don't have infinite energies. So that means that you can acquire this energy, and by acquiring more energy you acquire more voting rights, essentially, for voting on what is the main chain. So you vote with your energy.

In fact, there are crazy analogies where in proof of stake we use security deposits. So security deposits — Actually, before I bridge the analogies, I'll explain. So that's proof of work as you described. Proof of stake, what we essentially do is we have these bonds, and these bonds are locked up Eth, and that is you know the cryptocurrency in Ethereum, or A, the underlying one.

Now, thee bonds represent your voting rights, your weight. So the more Eth you have locked up in this smart contract, in this Casper proof of stake smart contract, the more voting rights you have for what is the actual main chain. So what we do is we are actually — We're doing this hybrid Casper approach. So instead of replacing proof of work full steam ahead in one fell swoop, we are doing a check pointing scheme. So that means approximately every 50 blocks we have a new checkpoint, and this checkpoint essentially is a single block, which validators, this is the people who have this bonded stake, vote on a particular checkpoint. This is a particular block at a certain height. Really, these checkpoints, the actual definition of them is just block number mod 50, and if that equals zero, then you're at the checkpoint. So you just vote on what that block is, and if you get more than 66% of the votes, then that is considered finalized, and what that means is that is in the main chain.

Now, here is the classic problem. In proof of work, if you were to vote on a single block, that means that you have to burn energy. So that costs you money. It costs you money to vote. Now, but in proof of stake, it doesn't cost you money to vote. It's just essentially signing a signature saying, "I vote on this particular chain."

So what has classically plagued proof of stake systems is this nothing at stake problem. The nothing at stake problem is basically what I just described, where in proof of work you don't have nothing at stake, because you just created a block that means you burned some amount of electricity. In proof of stake, you created a block, but that just means that you signed a signature. So what we have to do is we have to /validators, a.k.a. remove their bonds, remove their funds, delete their fund if we detect that they vote on two different chains, because remember, our goal is to create a single chain with a single history that is hard to revert and uncensorable.

So, validators, if they vote on one chain, we are able to detect essentially, and I can go into the mechanisms there, but we're able to detect if they voted instead of just on one chain, but on a conflicting chain, on another chain that is essentially a different version of the history. So this is — So it's a kind of bring it all together. We're trying to create one chain, and validators deposits some stake, that is locked up Eth, that is now used as their vote. When they vote on a particular chain, that constrains them to only that chain, and if they were to vote on the different chains, they would lose all of their deposit. So this mechanism, hopefully, you can start to see, this

mechanism basically says, “Okay. Let's say you vote. You’re voting. You’re voting. You’re voting over and over on block after block, you are able to form a single chain that is very difficult to revert.”

[0:23:51.4] JM: Before the check pointing, you've got these blocks that are accumulating. Are you seeing — Are there disjoint chains that are developing, and then you have a different chains that are being voted on at that checkpoint, or do the different validators have some system of keeping their chains aligned?

[0:24:12.4] CF: Yes. Good question. So, I mentioned this briefly, but the underlying — This is a hybrid Casper, and so we are actually still using proof of work on the underlying system. Proof of work is the block proposal mechanism. So proof of work is block proposal and proof of stake is like finality. The block proposal mechanism is exactly doing what you said. So proof of work, you’re burning electricity to create a block, which means that miners have this incentive to build on the longest chain, because then it's a higher chance that their block is included. So you get this, you build up the single chain, and then validators are confirming that chain, confirming that proof of work chain.

The validators, basically, what they do is they wait until there is a certain block height; 10, 20 blocks of proof of work blocks, and that gives them enough certainty to say, “Okay. Actually, I believe that this block, this particular block is going to be in the main chain.” So maybe block 100 is created, you don't know which one, if that's going to be in the main chain, because it hasn't reached the kind of depth, and you get this exponentially more difficult to revert property that that proof of work gives you. You wait until 10 blocks later, and now the current head is 110 and all the Casper validators start voting on block 100. They say, “Okay. this is probably in the main chain,” and then that finalizes the block.

[0:25:38.7] JM: So undercurrent proof of work, we know that this is a pretty sturdy form of coming to consensus on the transaction history, and I understand that the model for deploying Casper, which is Ethereum’s proof of stake, is it's going to be a series of different consensus mechanisms to rollout to full proof of stake. So we’re easing into it and we’re doing partial proof of work, partial proof of stake with this Casper-friendly finality gadget where you're saying, “We’re going to use proof of work for proposing blocks, and then we’re going to proof of stake to

finalize blocks. But if we've already got a proof of work system that works and we're not throwing it away, why is finality important? Are we giving up some degree of consistency or some degree of consensus in our proof of work? Are we sacrificing some consensus reconciliation with the classic proof of work and we're moving that to the proof of stake periodic check pointing?

[0:26:50.8] CF: So we're not really giving anything up. We probably will reduce the block rewards for proof of work, which means that, economically, it's pretty reasonable to guess that there will be less total difficulty on the chain, a.k.a., less people doing proof of work. But the real reason why we are moving to proof of stake — And, by the way, block proposal is done with proof of work in hybrid Casper, but full Casper is absolutely coming and being worked on and that completely strips out proof of work completely.

So the reason why we're moving away from proof of work, there are a whole slew of reasons. However, the first one and most clear cut is proof of work currently burns an obscene amount of energy. From one perspective, this is a practical; let's not destroy the earth with our consensus protocol. Just to give you an idea of how much, currently, the Bitcoin blockchain burns more energy than 156 countries, essentially. So this is a massive amount of energy waste.

The second reason is because in proof of work you actually get a number of undesirable qualities. Essentially, if you're able to 51% attack a proof of work chain, then you can do this spawn camping attack where you produce blocks, and there's no way to really remove your massive mining equipment, for instance, without hard forking to a different proof of work protocols.

If you develop a new ASIC, then it can be problematic. But that's actually not the only thing. There is also, in the USAF, soft fork. So there was this whole, essentially, user-activated soft fork and that created this incentive for the soft fork chain to arise, and this gets really complicated. But, essentially, the software chain, if it was popular enough, could have reverted what was at the time considered the main chain. So what this did is this proposed a systemic threat to the main chain. Like most of what crypto- economics in these protocols is doing is they are planning for the worst-case scenario. So they're essentially like minimizing that long to risk.

When I talk about these things, it may sound, “Oh! That sounds like a crazy thing. Why would that ever happen?” That's what we're trying to plan for and that's why you trust the protocol, is because we have a plan for each one of these terrible circumstances. This is really critical, because you could use a Google. You don't need proof of work. You can just have proof of authority. Google will sign off on your transactions. But then, once again, you have this risk.

The final thing that is actually very practical, is that once we moved to sharding, which we can totally talk about, which is scaling the blockchain, it's very useful to have an active validator set. So to scale a blockchain — Maybe I'll get into this later. But having an active validator set is really important, because it allows you to randomly sample from the validators and basically assign validators to shards. That is a critical security property, because if you are an attacker, you could focus your mining power on a particular shard otherwise and that would totally break the security of the system, because we want one secure mechanism, one secure consensus mechanism that then is able to propagate that trust throughout many, many different kind of chains or a large number of transactions on a blockchain.

[0:30:20.7] JM: Yes. Okay. I think I am seeing where we’re going with this. So we've seen that proof of work gives us a sense of security, and a sense of safety and, sense of liveness so far over the last decade and that makes us pretty comfortable with proof of work, but people are starting to pay more attention to blockchains and more attack vectors are going to be tried. It's worth building more insurance in, and the 51% attack, which we remain vulnerable to because of the miners got overtaken by 51% of some centralized evil, malicious force, then we would have a severe problem.

So we can create this proof of steak validation check pointing, which does not necessarily have an effect on the liveness, but it improves safety, because you have validator's that may or may not have mining power, but they have a stake in the system. They've got a bunch of money that they've put up in their bond. They’ve put up some money and said, “I am willing to bet that I am going to tell the truth, and if I don't tell the truth, I’m going to lose all my money,” and this is potentially even more useful in the future, because you could move to full proof of stake and you could just have these validators be doing the job of the miners and you would be able to parallelize the confirmation of transactions.

But just to go a —To keep going slowly. Am I correct about the fact that we're basically just adding a little degree of safety and we're shifting some of the control of the blockchain to people who have a stake in addition to people who still have that mining capacity?

[0:32:27.4] CF: Yes. Your description was very elegant, and basically — And I would say that, generally speaking, yes. We are getting to a — We are adding a level of security that doesn't exist today that will be critical for the proper functioning of a massively scalable blockchain. Not only that. I do, actually. One thing that you did leave out was that it is going to be burning a massive — Much, much smaller amount of energy.

Okay. I would say —

[0:33:00.4] JM: When you get to sharding.

[0:33:01.8] CF: Yeah. When you get to proof of stake, in general. Proof of stake is great for — I mean, it reduces the amount that you have to pay miners if you —

[0:33:10.9] JM: Even under the friendly finality gadget? Even under the partial —

[0:33:13.8] CF: Yes, because you're able to like reduce the rewards, and so that reduces the incentive for people to spin up new mining nodes. Yeah. Even under this partial model, you're still getting energy savings. How much is a question of macroeconomic.

But if you think about the total vision of Ethereum, you wouldn't be able to — The actual architecture is a lot less clean and clear if you are using a proof of work system. The fact that you get this validator set and you can use this bonded Eth to not only incentivize the creation of a single main chain, but you can also incentivize the validation of shards or the of shards. It's like, just from a design principle standpoint, this is like what makes sense. This is the natural construction for a blockchain of this type.

Proof of work is super elegant and super reasonable and is useful in a lot of different scenarios. But it is not the optimal kind of long term solution for the kind of blockchain that Ethereum and wants to be. And the kind of blockchain that Ethereum wants to be is a blockchain which

provides you trusted execution of smart contracts that are scalable and can power a large percentage of the economic activity on planet Earth, whatever.

[0:34:38.9] JM: Okay. So let's say the United States government steps in and launches a simple attack, where they spin up a bunch of accounts on Ethereum and they purchase 75% of the Ethereum in circulation and they use that 75% to have their stake. They make these big bonds and they spin up a bunch of validator nodes. Is the Ethereum network vulnerable to that circumstance where you have 25% of the wealth of Ethereum in decentralized account, like you and me, and 75% is the US government teamed up with Google somehow? Is that problematic?

[0:35:22.8] CF: So, this is an amazing question. The critical thing when you're designing these protocols is you want to constrain power. So what that means is that any constraint you can provide, some of these constraints aren't necessarily — Some of these are like fundamental constraints. So even if you do have 75% of the stake, let's think about what you can actually do.

So here are some of the actions, and this is how you would design any of these protocols, and I'm talking about Casper, but I'm also talking about any blockchain application built on Ethereum or another blockchain. You think about, “Okay. The central authority, what can they do?” Okay. One thing they can do is they can revert history. So they can cause a safety failure. That means that they voted on two different chains and those two chains both got finalized. So that means users are basically — Let's say you withdrew your money from a deposit on one chain, but it didn't actually go through on the other chain and some kind of accounting got messed up and you lost your funds. That's really bad. That's catastrophic failure.

Now, in that case, you will see at least one-third of all of the total deposits of the validators get slashed. The US government, in this case, is going to lose a large sum of money. So that is one thing that they can do, and that also has other ripple effects. Maybe that decreases the value of Eth in general. I'm not going to be able to give you a full recount of like what is going to happen, but this is — These are the actions in terms of these are the action.

So, now, the second thing is that they can start censoring. They can, for instance, censor other validators. Now, one cool thing about Casper in the way that we have designed it is that it actually has a censorship resistance built in. So how this works is if validators detect

censorship, then what they can do is they can actually create a soft fork. So they can start mining their own chain, a chain that is not censored, because the censoring party is on the main chain.

Now, what this soft fork can do is automatically, without even a hard fork, without having to cause a big social change where everyone updates their clients to remove the US government's coins. Validators on the censored chain will actually be able to — They lose money. So everyone's losing money in this — Well, not everyone, but the US government, let's say they have in a 75% of stake. So their chain is working seemingly normally, but the censored chain is not, from the censor chain's perspective. It's a little hard to get your head wrapped around this stuff.

But now from the censored chain's perspective, the validator's who are online — So the censored ones that we're leaving, they are going to be losing money, but they're losing money at a much less rapid rate than the off-line validators, a.k.a. the US government. So the US government is losing money very rapidly and the online validators from that chain, the censored ones, they are gaining money in proportion.

Eventually, what will happen, is that 25% of censored validators is going to become 66% eventually, of the total bond at stake. At that point, there is the potential for two chains to be finalized. Now, what that means is this is kind of a leaking mechanism, where a censoring majority is actually leaked out of the validation role. The reason why this is important, for a number of things, is what we're doing is where analyzing the incentives for all of the different parties involved. So one incentive that's actually pretty cool is if you believe that the censored chain is more — Or will win, and one chain is more valuable than the — That one chain is going to have all the value, then there's actually this incentive for you to defect from the censoring chain, the government's own blockchain.

So even in the case where — So all of these combined, even in the case where the government is going to be, let's say, reverting finality or it's going to be censoring other validators, even in these cases, they are constrained in what they can do. They have these disincentives. They have these mechanisms, these checks and balances which keep them in line, and that is an incredibly powerful thing, not only because in the case of catastrophic failure we have a plan of

action, but also because it discourages that catastrophic failure from ever happening, and that is what we're trying to do when we build these blockchains. We are trying to reduce the risk of catastrophic failure.

[SPONSOR MESSAGE]

[0:40:01.1] JM: Users have come to expect real-time. They crave alerts that their payment is received. They crave little cars zooming around on the map. They crave locking their doors at home when they're not at home. There's no need to reinvent the wheel when it comes to making your app real-time. PubNub makes it simple, enabling you to build immersive and interactive experiences on the web, on mobile phones, embedded in the hardware and any other device connected to the internet.

With powerful APIs and a robust global infrastructure, you can stream geo-location data, you can send chat messages, you can turn on sprinklers, or you can rock your baby's crib when they start crying. PubNnub literally powers IoT cribs.

70 SDKs for web, mobile, IoT, and more means that you can start streaming data in real-time without a ton of compatibility headaches, and no need to build your own SDKs from scratch. Lastly, PubNub includes a ton of other real-time features beyond real-time messaging, like presence for online or offline detection, and Access manager to thwart trolls and hackers.

Go to pubnub.com/sedaily to get started. They offer a generous sandbox tier that’s free forever until your app takes off, that is. Pubnub.com/sedaily. That's pubnub.com/sedaily. Thank you, PubNub for being a sponsor of Software Engineering Daily.

[INTERVIEW CONTINUED]

[0:41:45.2] JM: To put it in terms of safety and liveness, in this kind of events, if the United States government did decide to launch this sort of attack, there would be a period of time where the Ethereum network might not work so well. So you would have some liveness issues, but safety is guaranteed.

[0:42:08.8] CF: Yes, in the censorship case. In the case where they revert finality or they'd finalize two checkpoints, I should say, that is — Safety is not necessarily guaranteed, because safety basically means, "Okay. Is history going to change under your feet?

[0:42:24.4] JM: I see. Okay. Got it. Well, I think we've given up pretty good overview of some of the mechanics of proof of work and proof of steak as best as we can do on a podcast. A few more element to this, actually. So if I'm a miner, if I'm an Ethereum miner, how does Casper affect my business?

[0:42:44.6] CF: Well, it won't affect it too much, maybe. I mean, there may be reduced rewards. In the long term, you should probably create a staking pool or take your funds that you've made from mining and deposit them as Ethereum cerium validators. It is a pivot, but you can put your mining hash power on something else, mine Zcash or some other chain.

[0:43:09.8] JM: Or start a cloud provider. Why not?

[0:43:11.4] CF: 100.

[0:43:12.8] JM: Yeah. So what's the rollout plan for Casper? Because I know we want to talk about sharding and we want to talk about how he eventually get the full-on proof of stake. So we talked about this proof of work with proof of stake. How do you get to the full-on Casper proof of stake deployed for the entire network?

[0:43:32.5] CF: That's a good question. So, essentially, we're starting out. We just released a test net, and in fact there are now multiple clients that have the hybrid Casper code implemented, which is super exciting. So then the first step is getting hybrid Casper deployed on to the main net, and eventually that will require a hard fork. The general idea is that you can then reduce the mining rewards over time and increase the staking rewards over time and get to a point where you finally replace the block proposal mechanism entirely, and that is with one possibility, which is reasonable, is essentially these votes, they become blocks and maybe they become blocks on different shards. This is the area of more specification and research and not 100% strictly defined, like the hybrid Casper, but it is something that we were kind of working on as we also work on sharding.

[0:44:28.6] JM: Okay. Well, let's get into sharding. So we've talked about the resilience benefits of building a blockchain with proof of stake. What are the upsides? How is this going to move us to a world from AOL to — Well, I guess, I should say from dial-up to broadband, from default slow to default fast and performant?

[0:44:55.0] CF: Yeah. So, essentially, when we have this validator set, and even before then we can — What we're going to do is we're going to begin the — Well, we're already working on a sharding client, a sharded Ethereum client. So, the general idea, is you have this validator set and the validator set is randomly sampled and assigned — Each validator is assigned a block on a new shard. Now, the key behind sharding is, what we want to do is we want to scale the blockchain while preserving decentralization.

So what does that mean? That means that you can scale a blockchain a number of ways. You can scale a blockchain by creating, making huge blocks that are very computationally difficult to compute, and that only can be done by an AWS server or a quantum computer, or whatever. That is scaling, but not maintaining decentralization.

The way we need to scale, if we want to fulfill this kind of Web3 crypto-economic vision, is we need to scale while still allowing every laptop computer the ability to validate the network. We want to create a system where all the participants are small, relatively speaking, week computers, but that combined creates this resilient and secure blockchain, and scalable blockchain.

So the way we're doing is we say, "Okay. Validators, you don't actually have to download all of the blocks from all of these different shards. So think about each shard is its own blockchain, and validators are building blocks on those blockchains. Now, we don't want the validator is to hold all the blockchains on the computer because, remember, Ethereum currently runs on a laptop computer. If you were to run 100 Ethereum side-by-side on your laptop computer, your laptop computer would just explode. Not really.

The way that it works is these validators are able to be assigned a particular blockchain, they validate to check to see if it is valid, if those blocks are available. Once they see that the blocks

are available, if they are available, they build on the longest available chain, the longest — It's basically a chain that they can download.

So what we're actually doing is we're separating out this this validation from the — Like block proposals. So you'll now also have — You have two classes of citizens. You have this validator that's checking the blocks are available, and then you have this block proposer, which is downloading all of a single chain, which is basically what an Ethereum client does today, and proposing new blocks on that chain and executing transactions, basically, bundling up transactions into blocks, and validator's are saying, "Okay. Is that block available? Can I download all of it? Does it make sense? Let me include it in the — What I believe is the next step for this particular shard."

So what that does is that means, "Okay. Now we're able to build 100 different shards," which is what we're starting out with the 100 different shards while still maintaining this property that every validator, every participant is running on a laptop.

[0:48:07.0] JM: Just to drive the point home a little further for the efficiencies of a sharded blockchain. In today's blockchain validation mechanism, you've got miners racing to validate similar blocks of transactions. So each of them has access to a mem pool. The mem pool is roughly the same, and they're probably going to pull out similar transactions from the mem pool and decide which of those transactions go into a block, and all of the miners are competing to solve not exactly the same puzzle, but a very similar puzzle, and they're racing to validate overlapping transactions. Basically, out of N-miners, N-1 are going to have their work, essentially, go to waste. I mean, it's arguably not going to waste, because this is the only system that we know from empirical data, works to keep a blockchain consented upon in a safe way, but that's just empirical data, which is why going to prove of stake makes a whole lot of sense to me, at least trying it out, because the only way to get empirical data is to run the experiment.

Just emphasize — Well, I guess, talk a little bit more about it. Because with proof of stake you don't have to have everybody doing duplicate work, you can actually divide up the work safely. Really emphasize why this is important. Why is it so important that with proof of stake we are dividing up the mem pool and we're not having this duplicated work? How much savings for time and energy are we getting here?

[0:49:52.4] CF: So, essentially, the current Ethereum blockchain runs around 15 transactions per second. Now, with sharding, eventually, we will be able to run thousands of transactions per second, 10,000 more, whatever. Now, with this kind of version one sharding, you get, let's say, like a hundred axing trees. You have 100 shards. I mean, is not exactly this. However, you basically get these, like 100 parallel Ethereum chains that we know today, running side-by-side, all progressing in unison, essentially.

I mean, basically the biggest bottleneck right now with blockchains, the biggest thing holding the space back is the fact that we have the spotlight on us, and there are tons of developers that want to build the next Facebook, want to build the next Uber, want to build the next whatever, and they currently can't do that, and that is because there just isn't enough space in the blockchain. There isn't enough transactions per second. It can't handle it.

So what we're doing is were saying, "Okay. We need to make sure that all these applications that need to exist for this decentralized vision are compatible on the blockchain." So what we're doing is we're creating — We're working on sharding, which is going to speed up the transactions per second massively for the core Ethereum, and then we're also innovating on design practices so you can actually bundle transactions together, for instance, and submit less information to the main chain, but get more effect.

So Ethereum currently transacts about as many supports, as about as many transactions as Uber does. So if you were to implement Uber in a naïve way, maybe it would just fill up a single Ethereum blockchain. So what we want to do is we don't want to just support Uber. We want to support Uber, Facebook, and every competitor, the universe. We need to, instead of just having one Ethereum, we need 100 Ethereums, or maybe a thousand Etehreums, a hundred-X, a thousand-X, what we're doing today.

[0:51:57.3] JM: Okay. So, yes, I'm with you, like scaling transaction throughput is going to be great for scaling Ethereum. I don't think we have time to talk about plasma, unfortunately. But plasma is a way for scaling smart contract execution and it has something to do with sort of turning every smart contract into its own scalable entity. Maybe we'll have to do another show on that. But is the problem only transaction throughput or is it also a lack of other infrastructure?

Because, for example, Ethereum is the world's computer, but it's not the — It wouldn't make sense to do S3 with the Ethereum blockchain. Like you would probably do that on something like IPFS, and I have not really heard anybody proposing some sort of like data warehousing solution built on the blockchain or some distributed streaming service.

There's so much infrastructure that, for example, an Uber has built and that a Facebook has built internally, and maybe you could have some mixture of centralized and decentralized infrastructure, but I think what I'm trying to get at is like what are the other bottlenecks to getting to the decentralized Uber? Do we need IPFS to be in production and to be proven to work, or do we only need — Do we just need Ethereum and we can like to have — We can build the decentralized Uber while using S3 plus Ethereum, and then eventually migrate S3 to IPFS, for example. What's your roadmap for when we start to see people have these daps that are widely used by consumers? When are we going to start to see those?

[0:53:48.7] CF: Great question. So, to address, there are a number of kind of decentralized X or decentralized Y. For instance, decentralized file storage. That is something that IPFS, Swarm, Filecoin, Storj. There are projects that are working on this that allow you to essentially store a file and pay some crypto, and magically that file will still be around when you want to retrieve it a year later. This is an area of research that — And something that is actually in some ways enabled by the existence of an Ethereum type system.

So if you think about a lot of — When you actually said that some of these things can be centralized, this is, in some ways, correct. Where you can provide — The essence of what we're trying to do is we're trying to constrain central authority. We're trying to constrain power. So, for instance — And we aren't going to get into plasma. But just as a quick kind of example, the plasma MVP is actually very interesting, and that it has a central authority. It has central actor, but this actor is constrained and that they can't actually steal any money. So, essentially, people are able to use plasma to send hundreds of transactions per second to one another and still, that central authority, if it were to shut down or try to steal money, cannot actually do anything. It won't succeed in taking anyone's money. The money that is on the plasma changes leaves the plasma chain and goes back to the main Ethereum chain.

So the application stack that you're describing is going to be a product of the creation of this decentralized file storage. It's going to be the creation of better programming languages for smart contracts. It will be zk-SNARKs Starks and cryptographic primitives that we can use, accumulators, and it will be also like the design principles of where you put your authority, who you trust and how much you trust them, and in what ways you constrain them, and how that defines the way you trust them.

So this is — Maybe even an example of a decentralizing factor is the moving towards these client-side web apps in an almost mundane way where you have the application is just running on your laptop instead of running on a central server and spitting out HTML. So all of these pieces come together to form the decentralized web, and that is what we're going to see and what we're starting to see the formation of, but it does require a huge amount of work from developers, from tool creators, from even architects to pioneer ways to build these systems, and it's going to be a big effort. All those things that I listed, those are things that need active contributors. So please contribute.

[0:56:47.4] JM: All right. Well, I think that's a good place to wind down the conversation. Carl, it's really great talking to you. We touched on a lot of different things. I think we'll have to do another show in the not-too-distant future [inaudible 0:56:59.1].

[0:56:59.3] CF: Sounds good.

[0:57:00.3] JM: I've got all these notes on plasma that I didn't ask you about. Yeah, all right. Well, man, it's always a pleasure, and I will talk to you soon. I'll see you in New York, hopefully.

[0:57:08.2] CF: Yes, that would be awesome. Thank you so much for doing this, so much fun.

[END OF INTERVIEW]

[0:57:15.4] JM: Your enterprise produces lots of data, but you aren't capturing as much as you would like. You aren't storing in the right place and you don't have the proper tools to run complex queries against your data.

MapR is a converged data platform that runs across any cloud. MapR provides storage, analytics and machine learning engines. Use the MapR operational database and event streams to capture your data. Use the MapR analytics and machine learning engines to analyze your data in batch or interactively across any cloud, on-premise or at the edge.

MapR's technology is trusted by major industries like Audi, which uses MapR to accelerate deep learning in autonomous driving applications. MapR also powers Aadhaar, the world's largest biometric database, which adds 20 million biometrics per day.

To learn more about how MapR can solve problems for your enterprise, go to softwareengineeringdaily.com/mapr to find whitepapers, videos and ebooks. MapR can leverage the high volumes of data produced within your company, whether you're an oil company like Anadarko or a major fin-tech provider, like Cabbage, who uses MapR to automate loan risk and has done $3 billion of automated loans to date.

Go to softwareengineeringdaily.com/mapr to find out how MapR can help your business take full advantage of its data. Thanks to MapR for being a sponsor of Software Engineering Daily.

[END]