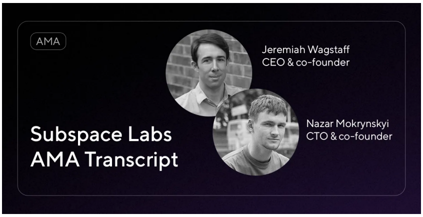
Subspace Labs AMA Transcript | February 2022 - Subspace Network

By Subspace Network

Source: https://blog.subspace.network/subspace-labs-ama-transcript-february-2022-59edbc67ec35

Co-founders Jeremiah Wagstaff and Nazar Mokrynskyi held a live 60-minute community AMA in our Discord to answer questions about the Subspace Network.



For the recording of the AMA, please visit the Subspace Network official YouTube channel.

Introduction to the Subspace Network

Jeremiah Wagstaff: My name is Jeremiah. I am the CEO and co-founder of Subspace Labs. In case you are not aware, we are building the Subspace Network blockchain. Subspace Labs currently has a remote team spread all over the world, across 7 different countries. We have 12 full time team members. We've been around for a few years. This is our fourth year working on the project. We have primarily been a research focused project for quite a while until recently when we started building the blockchain.

We've been obsessively focused on solving some of the core systemic problems in the industry. Things like energy efficiency, environmental sustainability, how to operate at global scale in order to support mass adoption, and how to handle the increase of centralization we're seeing in the industry.

So these are the things that really get us up in the morning that we really care deeply about. So we're building the Subspace blockchain to solve these problems. Subspace is a new layer one blockchain. We consider it to be the first fourth-generation or gen four blockchain.

What we mean by that is: first-generation obviously Bitcoin. Ethereum was really the next major upgrade, as the next improvement in the idea of blockchain. There are many competing third-generation protocols. You've got Polkadot, Near, Solana, Cardano, and Ethereum2. All of these make a series of trade-offs in order to achieve scalability on this so-called blockchain trilemma.

We call ourselves a gen four chain because we fundamentally overcome this trilemma. We don't have to make trade-offs in order to achieve scalability. And so we are in our minds and we look very closely at what other people are doing in the industry. We follow it very closely. We believe we're really the first blockchain that can really achieve global scalability.

What we mean by that is we can support mass participation in consensus, by ordinary people with commodity hardware. So this is kind of back to the ideal of single computer mining that Bitcoin started with, but we're doing this without the waste, without the electricity cost of doing this, we can support mass adoption on the user side so we can have billions of accounts, billions of transactions, millions of transactions per second, actually. And we can have blockchain-based decentralized applications that can actually run at internet scale. So Subspace as a blockchain, it's really just a platform for both storage and compute, meaning smart contracts, that is just fundamentally a new level of scalability.

The kinds of things that we think the platform will be used for, the kinds of things that we're really excited about, and that our chain is really, really well suited towards — probably are going to be built around NFTs. The ability to not only mint an NFT perhaps on Subspace, perhaps on some other chain, but also to store the data associated with that NFT.

The data doesn't have to be stored on a centralized server or controlled by some company that issues the NFT, so that data has the same permanence as the NFT itself. And then the things around that with gaming and the metaverse, and many of these other NFT use cases, which ultimately just boil down to billions and trillions of NFTs one day, all with lots of data associated with them.

So that's what we think is really going to be one of the killer apps or key use cases of Subspace in the future, but it's really a generic and abstract platform at the end of the day. So that's kind of a high level about Subspace. I'm going to turn it over to Nazar and let him talk a little about why he's excited about Subspace.

Nazar Mokrynskyi: Absolutely. Thanks for coming guys. I am Nazar. I'm also a co-founder and CTO at Subspace Labs. I think one of the most exciting things about Subspace, and I'm biased of course, is that in contrast to many other projects, we fundamentally do not accept the blockchain trilemma. We do really think that you do not necessarily have to make trade-offs between scalability, decentralization, and security of the network. We can kind of have all that in one place. And in order to achieve that, you need to really go into the fundamentals and start from the first principles to design a network, which has a fair consensus, which is based on the resource, which is widely distributed storage.

Ideally that resource shouldn't be wasted. That's why our consensus is actually proof of useful storage. We call it Proof-of-Archival-Storage (PoAS), where all of the blocks and transactions that people submit to the network will be archived by the farmers (that's what I expect most of you guys will be doing). And then, proportional to how much data you store of the history, like the useful service objectively, proportional to that will be your chance of winning into farmer block reward. And we believe that decentralization is very important and that is why all of the blocks that we have and all of the farmers that will be competing for the block rewards (all of those rewards will go to farmers) are a big part of the tokenomics of the project. We really believe that is a key part in contrast to some of the networks which are using Proof-of-Stake, which is basically a permissioned environment where you need to have coins to participate.

On the other side, storing data is fine, and having consensus is cool, but we also do want to have some computation on the blockchain. That's the big innovation that Ethereum brought to the table. And we have several parts to that. One is the decoupled execution where farmers produce blocks, but they do not run the blocks. That's why farming on Subspace is very energy efficient, and anyone can basically do it as soon as they have a little bit of space.

There is another role in the network, which is executors. They will be a bit more powerful machines like gaming desktop computers that can run transactions and earn rewards like fees from transaction execution. Those will also be distributed probably a little bit less than the farmers, but it's important to remember that even though those nodes are powerful, the control over the network is still in the hands of the farmers.

This is a big distinction between Proof-of-Space, Proof-of-Archival-Storage that we have, and Proof-of-Stake, where executors are kind of like validators in those networks. But the one that does control the network in our case — farmers control the network. All the executors just do is deterministically, run whatever farmers decide to do.

And one part of the protocol, which is not very well described right now on our website, but it is kind of mentioned on the technology page is the scalability. So our ability to scale vertically shows that executors can run basically as many transactions as they can afford in terms of CPU, bandwidth and other resources. But we also have an interesting design, which flows from these decoupling of farmers and executors, which allows us to have a pretty unique, horizontal scalability. And that is the part where if you have, let's say a thousand transactions per second, with one shard (which is not amazing by today's standards, there are networks that do more), but let's say that's a conservative number.

If you add a thousand shards, that's a million transactions per second. And our network is designed fundamentally to allow us to have as many shards as the number of executors on the network and still have decentralization in terms of consensus, because consensus is done globally, in contrast to some of the other networks where you have to split block producers between shards and throw out some secret duplications of that.

So I think this unique combination of both, sharding of execution of storage processing and scalability in terms of consensus participation — all of that together is a pretty unique combination that fundamentally overcomes the famous blockchain trilemma. And if you don't really believe that it is, it is true. It has to be that way.

What are some competitors that exist or are already in development currently?

Jeremiah Wagstaff: Depending on which aspect of Subspace you're talking about, there are definitely projects that are doing similar things. On the consensus side with what Nazar was talking about with Proof-of-Archival-Storage and farming, Chia and Filecoin would be two of the biggest competitors, I guess you could say, or maybe comps.

And the real difference is that we're really aiming to be as permissionless as possible. In Filecoin, you have to stake. You also have to have some very special hardware and it really only supports this storage use case. It doesn't have generic computation or scalability built into it.

In Chia, they're much closer to our ideal of permissionless consensus. But at the end of the day in practice, they burn through SSDs and they require a lot of energy in order to do the plotting, and they don't have storage. We've really carefully designed our plotting so that it's much more energy efficient and it works on any hardware. And Chia kind of supports improved Bitcoin-style, smart contracts, but it's not like a global computation model, like Ethereum. And again, it's not designed for scalability. They're just kind of both really fundamentally generation one blockchains that have a different consensus mechanism than Bitcoin. We have something which we get compared to on the storage side a lot which is Arweave. Nazar, I'll let you take that.

Nazar Mokrynskyi: Yeah, on the Arweave side, they did a great job in terms of offering the archival storage. But fundamentally they are also kind of like the first generation of storage networks, because of their architecture. It's not really possible to increase the throughput like we can with sharding. Basically, let's say if you have like 10 megabytes per second per shard and have a thousand shards, suddenly, we have almost 10 gigabytes per second of upload that we can add to the network. That will scale with the number of participants in the network instead of getting slower. You can consider Subspace as an evolution of that idea. Also importantly, in our view you have the process of mining. So you still burn CPU resources to participate. Storage is kind of a side feature for us. The storage is the primary mechanism. We don't have to burn electricity just to participate in block production. Your storage is your ticket.

Jeremiah Wagstaff: One other comment on Arweave is the pricing model. This is really one of the key innovations which we introduce. We have a dynamic cost of storage. So it responds to supply and demand. You can think of it like an automated market maker for on-chain storage. Whereas in Arweave and pretty much all blockchains where you're storing data directly on the blockchain, there's a fixed cost of storage. This is like Satoshis per byte in Bitcoin or like gwei per byte in Ethereum, and it doesn't really change. In our case, it changes based on how much supply and demand there is. We can actually measure the space pledged to the network as part of consensus — we can also measure the demand. That's just the size of the blockchain at the end of the day, because everything is just being written to the blockchain — all the data that's being stored on the network. So this is the most incentive-compatible and it also means that as we get more adoption on the farmer side, that the cost of storage actually gets lower and it becomes more attractive. It becomes more economical which just drives more adoption. So we get this kind of virtuous cycle that will eventually find a true cost of storage.

When do we plan on opening up the testnet to the community?

Jeremiah Wagstaff: Well, I would say technically it is already. We just haven't really been talking about it. So if you wanted to, you could go and just run the instructions that are in GitHub and you could run a node on the network, but just to be clear, it's not incentivized. So there's no reward for doing this. We're still working out bugs in the protocol.

We're still trying to improve the user experience. It's going to be breaking continuously, like probably on a weekly basis. We're still breaking our testnet as we upgrade and change it. So really anybody is welcome to participate and we would definitely welcome the feedback. If there's anybody that's like a tester out there that loves to play with these new protocols and give feedback, we would welcome that.

But we're not really going to be announcing community farming and trying to build up the community probably, you know, for at least for a few weeks. But our next major objective is going to be launching an incentivized testnet. We still don't have a firm date for that yet, but we're trying to get it out as soon as we can.

Nazar Mokrynskyi: Yeah, all of our team internally is already running a farmer and node implementation, and we are testing how things work. I know some people from the community already found the repository and just went through instructions and successfully joined the network. So that was awesome to see.

Discord User: Hi guys. I just have a quick question about the NFTs. I'm not very tech savvy, so the answer may be obvious, but just bear with me, so let's say I mint an NFT on Solana, Avalanche, or I don't know, some protocol. Why would I store it on Subspace instead of the original chain? I mean, what is the advantage?

Jeremiah Wagstaff: Yeah, that's a good question. Thank you for asking because it's not always obvious to people. So there's really two parts of an NFT. There's some ownership data about the NFT which has to be tracked on the blockchain. That's the actual asset. And then there's some metadata or some content data that's often associated with it.

So if it's a PFP, there's an image, or if it's an audio file like music, there could be an MP3 or it could even be a video. If it's a metaverse object, it'd be something like an asset file. Blockchains have a really, really, expensive bottom line to store this data. It's like a hundred thousand dollars per megabyte or something like that on Ethereum. To store data, it's cheaper on other blockchains, but even those are really not optimized for storing that data. And they tend to not be stored on the blockchain and they tend to be stored on some centralized provider, maybe hosted on IPFS, but not guaranteed to be on IPFS.

So we're still saying put the asset data on that chain but put that file data on subspace and give it the same permanence. So that if the business shuts down or if the IPFS node gets turned off, that data is still available.

It's also about security. It's double security just in case the actual thing gets destroyed. In the original your data still exists. As long as the blockchain exists, that data will exist.

What is the advantage of splitting executors and farmers?

Nazar Mokrynskyi: That's actually one of the components that allows us to have a high decentralization on the consensus side. So imagine when you have farmers which not only have to produce blocks, they also need to run all of the transactions and computation. This is basically what happens in Ethereum. Then every node who is a farmer would have to have the hardware capabilities throughout all of the transactions, or in other words, the throughput of that blockchain will be limited by the slowest computer that is on the network.

That is why in Ethereum, for instance, they cannot just crank the transaction throughput or the gas limit on every block. They need to target a specific machine. Because of that, they are limiting the amount of mutation that can happen every block.

Once we have a split, we segregate the responsibilities. We can have very lightweight farmers which can participate in consensus and offer security to the network, but you can also have more powerful machines. And those machines that do computations, they are not responsible for security. They are only responsible for computation. So, that is the approach to allow more work to be done every block without overwhelming regular farmers.

What happens to data with prohibited content that is hosted onto the Subspace Network, specifically illegal pornographic material. Will there be a policing feature to delete files from the blockchain? And if not, what are the legal implications of farmers running a farmer node?

Jeremiah Wagstaff: That's a great question. So we're really big believers in decentralization and being as permissionless as possible. Once you start trying to censor content, which you think is bad, or once you add the ability to censor content, then it can be used to censor other content because it's a very subjective question.

I'm not saying the stuff that you're proposing should be supported, it definitely shouldn't. But it becomes very subjective about what is and what is not censorable. So we are not planning to implement anything directly to control this. We're just building a reference implementation of this blockchain.

But it doesn't stop somebody from building another implementation, or extending, submitting a pull request. It's all open source. They can fork it, they can clone it, they can submit PRS or work on a substitute improvement proposal process. This is kind of up to the individual, how they want to handle that. On the legal side, I'm not a lawyer so I can't answer that. Of course it depends on what jurisdiction you're in, but it's definitely a regulatory question, which many people and many governments are grappling with right now.

Nazar Mokrynskyi: I can probably answer it from a slightly different angle. The way data is stored on the Subspace Network is you're not storing files directly.

The history of the blockchain is actually split into multiple chunks. Then you, as a farmer, store a random selection of those random chunks in a specially encoded way. So it's not like you are storing a particular file. It's more like the network as a whole stores some content.

Are there any hackathons or developer grant programs?

Jeremiah Wagstaff: Another great question. We do not have any hackathons planned right now. We've actually been attending some hackathons, building out some ideas we have, using the subspace API, and the JavaScript SDK that we have. And we do not have a grant program yet, but we will eventually. We will eventually have hackathons. I'm saying conservatively, maybe in the next six months we might have our first hackathon.

Grant programs will probably come closer to mainnet launch after we've done a community sale. That's where the funding is going to come for those grant programs. But it's definitely something that we have on our road map.

Where is the data stored?

Nazar Mokrynskyi: It is stored on the blockchain. It sounds ridiculous to some people that are coming from an Ethereum background, but basically what you do is you put your data into a transaction, submit a transaction to the network, and the transaction gets stored. Farmers do the work. And then that data is archived as part of the regular history.

Jeremiah Wagstaff: Just to expand on that, if we were just a single blockchain where everybody stored a full replica by default, that wouldn't scale very well, but there's a lot of our protocol that is designed around, basically at a high level, the way a data center works when it shards data across a bunch of different computers. We just do this over a network level and we do it with cryptographic incentives, but effectively, as Nazar said, the state gets chunked up. It gets erasure coded. Then there's a load balancing policy and a decentralized retrievability policy.

So all of this stuff can still be found and still coexist, even in these tiny little chunks that are spread to the four winds across the network. So any particular node that's storing the history doesn't really matter. They don't have to have as much as the whole blockchain history has. They just have whatever percentage they have. If they have enough to store 1% of the history, then they store 1% of the blockchain history, and then they get a percentage of the block rewards based on the total storage across the network. If they have more storage than the network (the blockchain), they could store multiple copies because they're storing unique copies. That's really an important part of our protocol. People are storing the canonical copy of the blockchain. They're storing the transformed copy of a permuted copy, which they prove they're doing. They're proving that they're storing it in this unique replicated fashion. This is how consensus works.

In that case, how do we avoid slow transactions that might occur? And how do we avoid data accumulation or a congested network.

Nazar Mokrynskyi: I think the answer to both of those is sharding. So when you have one chain, which is very congested, that basically means that there is more work to do than whatever one chain can do. And you can add another shard. So let's say we have one shard that can support X transactions per second, do X megabytes per second, etc. But the response is, we just add another shard and then we can split the load to half. And depending on how much demand there is and how many farmers and how many executors we have on the network, we can just keep adding those shards as needed, and that way we can scale.

What is our thought process with onboarding other blockchains? How do we convince them that using Subspace will improve their experience with their particular blockchain?

Jeremiah Wagstaff: That's a great question, and the answer is: we don't have to approach them. Meaning we don't need their permission. The beauty of blockchain is that this is all public data. It's all permissionless. We can kind of just integrate at will or even our community or developers could extend our protocol to integrate it with any other blockchains, kind of like with Uniswap. They just built a protocol and anybody can add a token pair to Uniswap.

That's kind of how we designed the way Subspace integrates with other blockchains. So we've built this for Kusama with the relayer that we have right now, kind of as an experiment to make sure that we've done it in a generic way. We've worked hand-in-hand with a lot of teams, and directly with Parity, and Web3 Foundation. Also with several of the parachains to get feedback on this, and everybody's been very supportive of this app, and Gavin Wood even talked about it in one of his last newsletters.

We're kind of taking the approach of doing this one ecosystem at a time. So KSM was the first ecosystem and DOT is going to be the next ecosystem. We definitely want to get buy-in and consensus from that community, specifically from the developer and protocol engineering community.

But we found that everybody's been, very, very welcoming and excited about these things, versus being hostile. They see it as kind of a positive thing because we're just helping them scale, helping them be a more decentralized protocol.

Will there be a cap supply for the token?

Jeremiah Wagstaff: Yes, there will be. The current tokenomics are effectively a Bitcoin style halving — just more frequently. Right now that's one year every time there will be a halving. Another important thing about the tokenomics I didn't mention is that over 51% of the token supply is going towards farmers, towards the community.

Executors are only getting the transaction fees. They're not getting the new tokens that are being created. This is all going to be part of our token white paper, which we hope to be releasing in the next few months, hopefully, but we'll see.

For farmers, is it better to use hard drives or SSDs?

Nazar Mokrynskyi: We made the farming process as lightweight as possible. So there is really no demand for powerful CPUs or fast SSDs. Regular hard drives should be just fine. If one has some capacity on an SSD, that's totally fine as well, but it might not be economically viable or as useful for that.

It's more targeted, like if you have an extra hard drive laying around. Let's say you have a hard drive for games, it's a few terabytes and you have half of that empty. You could just use that to farm some Subspace credits.

Justin Hill: To add to that, SSDs have a limited lifespan compared to a hard drive. While it is physical, it doesn't necessarily have a hard cap, like an SSD does.

Are there any requirements for the network?

Jeremiah Wagstaff: Yeah, that's a good question. So it goes back to what Nazar would say about having low bandwidth or low hardware requirements for farmers. So, in our protocol, we can really make this whatever we want, but the higher the bandwidth requirements are, the less decentralized it's going to be — more centralized, and harder to run a node. So we're going to try to keep this as low as possible. Effectively, we look at this as a community decision. So there might be a parameter that could be software worked or done. Since we're using Substrate this would be a forkless runtime upgrade to control this. Kind of like an Ethereum where the miners get to vote what the block gas limit is. That's more about the CPU bandwidth than the network bandwidth, although they're related for sure. We're basically going to keep this as low as possible. So even if you're in a country or a part of the world that has bad internet, you should still be able to participate in farming.

What is the Farmer’s Dilemma?

Nazar Mokrynskyi: The Farmer’s Dilemma is wild. The problems lay with some of the earlier Proof-of-Capacity limitations in general consensus. For instance, Chia is susceptible to this. The basic idea is that you have several things that you need to store for the blockchain. You have the history, you have the state, and if you have a Proof-of-Capacity blockchain, you also have the capacity that you provide in order to participate in block production and get rewards.

Essentially you have one disk and you need to store some space, so you occupy some space to participate in the consensus, but we need to also store the blockchain data. Since your award, as a farmer, depends on how much space you can provide to the consensus.

The logical choice is to run a light client instead of the full client, and to not store the whole history, the state, and they will likely just join a pool. This isn’t the greatest decentralization of the protocol, because then only farming pools control the network, which is very few compared to the number of farmers.

This means only light clients will be around the network for the majority of nodes. They will only store a small portion and they will be able to validate that things are correct. But there is really no guarantee that the history is recoverable.

And, basically this is called the Farmer’s Dilemma. Like what Farmers will prioritize, in terms of game theory, and by using storage, we just flipped this problem onto itself, where the thing that you are storing is the portion itself. So you don’t have to decide what you want to store, right?

You’re storing the blockchain and that is what is being used to participate in block production. This is like another small part toward the concept of decoupled execution, that executors will store the state like the balances of accounts. How much coin everyone has, that kind of thing. And the state on the farmer's side will be very, very small.

So there will be no incentive for farmers, or there will be very little incentive for farmers to not run a node. Because that's what would be like the default behavior for them. Hope that that makes sense.

Jeremiah Wagstaff: And another way to think about that is this problem of centralization in blockchains and economies of scale. What we've tried to do, and what we believe we've achieved is to remove any economies of scale from farming so that there's no incentive to pool your disks together, or even to go out and buy disks for that.

The only incentive would be just to reuse excess capacity that you already have. And one of the big lessons that we've learned since we released the white paper, is that this Farmer's Dilemma is actually just a special case or an instance of a much larger, much deeper problem, which is the problem of blockchain bloat. So as their state blows up, several hundred gigabytes in Ethereum right now that you need to maintain the state as the history blows up. I think it's nine terabytes to run an archival node on Ethereum.

This leads to increasing centralization on its own. Even though there's not a direct economic incentive, such as in Proof-of-Work or Proof-of-Stake, you're not penalized for storing this data. Whereas like in things such as Proof-of-Capacity, you actually penalize this action as there's a disincentive for storage that isn't being used for rewards.

But using your disk for this data, it still just becomes burdensome. And this is the problem of bloat, which leads to more centralization, again, and there's just a completely different security model.

Like all the things we talk about, like with Bitcoin, that'd be an honest majority assumption, or the safety of Bitcoin, the immutability of Bitcoin, these things don't work in a light client model.

It's a whole different security. You can double spend, you can create money out of thin air. There's all these things you can do if everybody's a light client. This is why people care so deeply, like the whole Bitcoin civil war was about this block scaling debate.

This is the whole reason that Ethereum has really resisted scalability on layer one and pushed the Layer 2's into the role of solutions is because they're concerned. And Vitalik is very outspoken about maintaining the ability to run a node on an average developer's laptop to run a full node.

When we solved this Farmer's Dilemma, we did it specifically for "how do we have a fair decentralized, Proof-of-Capacity consensus protocol, and more generally, how do we solve this problem with scalability as you scale?"

As you have thousands of transactions or millions of transactions per second, history in this state grows tremendously. And then nobody can run it or at least everybody has to run a lite client. So, we were able to take this work we did around our novel consensus and apply all of these amazing scalability techniques that other protocols have been working on, those that the broader blockchain ecosystem has worked on. Specifically, some leading researchers at Stanford that we've been working very closely with, and we can implement these ideas, and we can scale, and we can still maintain decentralization. We can resolve this problem of blockchain bloat.

Is it possible to determine exactly which specific chunks of data farmers are hosting? I think he’s referencing some of the Arweave farmers who were targeted by some nation states.

Nazar Mokrynskyi: Yeah. I think it’s not directly possible. There is logic that is in the implementation in which you only store data in small pieces. And we expect there will be alternative implementations eventually, but otherwise you are just storing four kilobyte chunks of random data in large quantities. You don’t really know what it is. It’s just random data for you until you know what it is and how to interpret it and how to assemble together. Until then it is just random bytes that someone may or may not store.

Jeremiah Wagstaff: To expand on that a little bit, there is a retrievability aspect where you can reconstruct objects or files that are put out onto the network. It is very, very highly distributed though.

It’s not about knocking out a farmer who has some file. It’s about knocking out hundreds or thousands of farmers who have all of the different chunks and you’d have to take out over 51% of the chunks, basically because of the way the erasure coding works. So, it is theoretically possible. Practically, it’s very, very hard to do, especially with the levels of decentralization that we expect to see, and the levels of adoption we expect to see in the farmer network.