Fidelity DIGITAL ASSETS℠

NOVEMBER 2023

Understanding Proof-of-Work

A nontechnical introduction to bitcoin mining and the implications of proof-of-work for investors

Chris Kuiper, **CFA**® Director of Research, Fidelity Digital Assets℠
Matt Hogan Research Analyst, Fidelity Digital Assets℠

Introduction

Investors are generally aware of Bitcoin and its related mining activity, which uses electricity. However, many may still be unfamiliar with why Bitcoin uses electricity and may question if it is useful or necessary, or if it could be replaced with something else. The purpose of this paper is to more fully explain bitcoin mining, or "proof-of-work," and to show how this process is linked to several properties of bitcoin.

It is vital for any investor considering bitcoin as an investment to understand at a deeper level what bitcoin mining is, how it works, what it accomplishes, how it may differ from other systems, and what it means for an investment.

The scope of this paper will be confined to understanding proof-of-work and, therefore, will not delve into the environmental debate of bitcoin mining, such as how much energy it uses, from what sources, and its emissions. Investors must first have a foundational knowledge of proof-of-work and the problem it solves to effectively inform further analysis regarding energy use.

Please also note that this paper favors analogies over more precise technical explanations for how proof-of-work operates and is not meant to be a technical reference for proof-of-work. Instead, the goal is to educate investors up to the technical level needed to assess properties of bitcoin as an investment.[1]

The structure of this paper is as follows: It starts with a nontechnical introduction to bitcoin mining, covers the governance or checks-and-balances of proof-of-work, and then the implications of the system and how it contributes to different properties of Bitcoin. Then, this paper will introduce and explain at a basic level the alternative consensus mechanism, proof-of-stake, and its implications. It concludes with comparing the two to better understand proof-of-work and why proof-of-work could be best suited for Bitcoin and its unique investment thesis.

Key Takeaways

- Proof-of-work is one type of consensus mechanism that allows a decentralized network, such as Bitcoin, to come to an agreement.
- Proof-of-work solved long-standing computer science problems and is key to the entire Bitcoin network's function.
- There are still some misunderstandings when it comes to how proof-of-work operates, leading to incorrect assumptions, which will be addressed.
- Proof-of-work's link to the physical world and use of real-world resources is one of its primary features and competitive advantages that lends value to the token and network, while also enhancing and solidifying its role as a digital commodity and emerging monetary good.

1 Proof-of-work can be and is used by other digital asset networks, in slightly different variations, but, for the purposes of this paper, the discussion is confined to proof-of-work as it is implemented by the Bitcoin network.

- In other words, the physical cost of computing and electricity is the feature of the proof-of-work system, not an undesired externality or byproduct.
- Proof-of-stake does not use the physical world layer, removing some natural or physical limitations and allowing virtual parameters to be more easily changed.
- Proof-of-stake and proof-of-work are similar in that they both require a cost of capital and both feature an incentive structure to keep actors honest, but in different ways.
- Compared with proof-of-work, proof-of-stake has different attack vectors and relies more heavily on governance and consensus at the social level.
- The proof-of-work consensus mechanism contributes to making bitcoin among the most—if not the most—secure, decentralized, and sound digital monetary goods.

Keeping the End Goal in Sight

Before getting into some of the more technical aspects, it is important to remember why Bitcoin uses a proof-of-work system and the role it plays in the Bitcoin network. In short, proof-of-work is what makes the Bitcoin network...work! It is a key piece of technology whose invention predates Bitcoin, but was implemented in a novel way with Bitcoin, along with other pieces of technology, to solve a fundamental problem in computer science.

The Bitcoin Network's Technological Breakthrough

At the most basic level, the Bitcoin network consists of many computers all running the Bitcoin software. These computers are connected to each other and make up a type of payment network that facilitates the transfer of a digital asset, also called bitcoin.

What makes the Bitcoin network different from all other payment networks is that it is decentralized—there is no one person, government, entity, organization, or corporation in control. This is the first of its kind in history.

Bitcoin was invented to solve the problem of creating a true peer-to-peer form of digital cash, where transactions could be done online with no central party or intermediary needed. Furthermore, the technological breakthrough also included solving the double-spend problem: making sure a virtual token cannot be spent twice or copied, giving rise to the possibility of true digital scarcity. How Bitcoin achieves this mechanism of fraud prevention is especially important because it is accomplished without reliance on a third party or intermediary. Up until Bitcoin's invention, this was something that had never been solved before.

This is done by creating a shared digital spreadsheet of sorts. A copy of this spreadsheet, or ledger, is distributed among all the computers running the Bitcoin software and each participant updates their copy of the ledger when new transactions are validated.

The question, of course, is how does one get a decentralized network of different participants to come to an agreement if nobody is in charge? Who has the correct or true copy of the ledger and how does one get everyone to agree on which copy is correct? The answer: a **consensus mechanism.**

Proof-of-Work: The Bitcoin Network's Consensus Mechanism

Ideally, a consensus mechanism should be reliable, fair, efficient, and transparent. The simplest way to come to a consensus is through a voting process.

Consensus as a Voting Mechanism

If voting is a way to achieve consensus among a group, it may be tempting to have a typical voting system of one vote per person for the Bitcoin network. However, this would face the challenges of the current political voting system of a central or third party needing to make sure each person only voted once.

What about one vote per computer connected to the Bitcoin network? Pseudonymous Bitcoin creator, Satoshi Nakamoto, also addressed this issue in the Bitcoin white paper:[2]

> *"If the majority were based on one-IP-address-one-vote, it could be subverted by anyone able to allocate many IPs."*

In other words, there would be nothing to stop a malicious actor from spinning up multiple computers or generating many IP addresses to outvote the network of honest actors.

But what if each vote had a cost? This is what proof-of-work achieves at a very simple level. As Satoshi stated: "Proof-of-work is essentially one-CPU-one-vote." By "CPU," Satoshi meant central processing unit, or a computer's main processing chip that performs calculations. In other words, to have a vote, you not only have to spend money to buy a computer chip, but you also have to run it and make it perform actions, costing electricity.

By making a vote have a cost, someone looking to tamper with the ledger's legitimate copy would have to endure a physical and, therefore, financial cost. As seen later, this aligns incentives because it would be more profitable to use these resources to support the network rather than to try to attack it.

> To summarize, at the most basic level, proof-of-work is a way to achieve consensus among a decentralized network with no single entity in control.

Next, this paper will go deeper and walk through the process of where proof-of-work fits into a bitcoin transaction and what the process of performing proof-of-work actually is. The purpose of the next section is to provide only enough of a technical explanation to illustrate concepts and properties of the Bitcoin network and token.

2 See https://bitcoin.org/bitcoin.pdf

Going Deeper on Proof-of-Work

The previous illustration of a voting system is helpful in thinking about how a decentralized group can come to a consensus, but it may not be the best illustration because it may lead to the belief that the Bitcoin network is a majority rules system. However, the Bitcoin network's ultimate goal is for the ledger to be true and the majority is not necessarily true or right. In other words, a majority of participants could still collude and change the ledger to something that is false.

Ideally, many would not want a majority in charge of the ledger, but rather one bookkeeper in charge who is always 100% honest and incorruptible. However, this idea once again introduces the following problems:

A. Having to trust one controlling person or entity.

B. How does one trust a person or entity without knowing something about them? This introduces the need for revealing personally identifiable information.

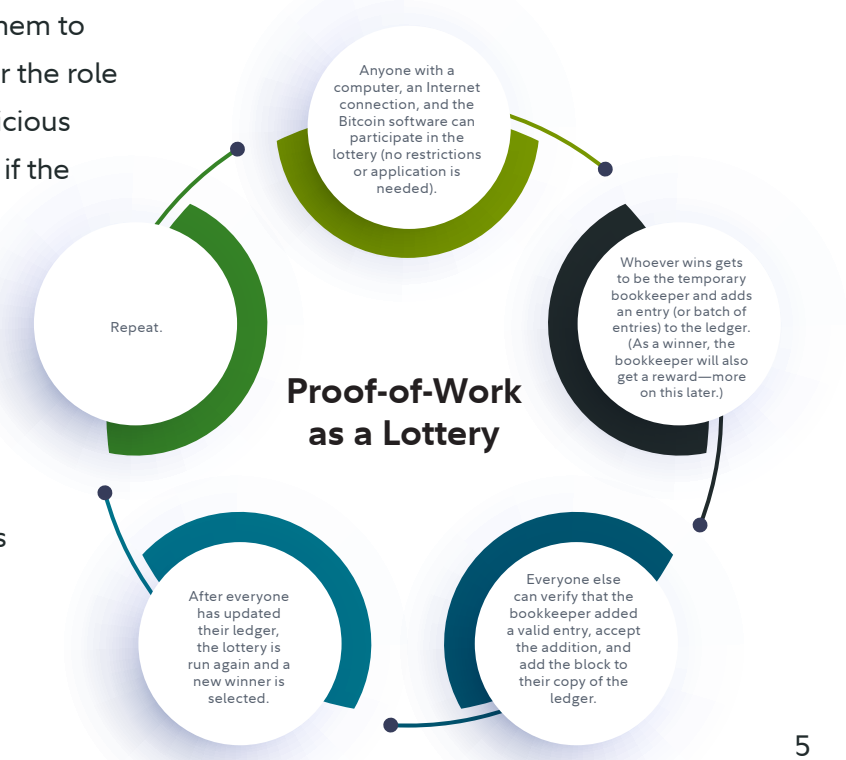
C. Once that information is known, a malicious actor could contact and bribe or coerce the person or entity to act dishonestly.

As Bitcoin educator Yan Pritzker elegantly points out, the solution to these problems is to have a kind of lottery that picks who gets to oversee the ledger and make changes to it. In his book, *Inventing Bitcoin*, Pritzker goes through the following example of how this could work and alleviate the problems above.[i]

Proof-of-Work as a Lottery

If one person or entity is chosen at random who then gets the privilege of adding to this shared ledger, then the problems listed above can be alleviated. Randomness solves the problem of knowing the entity ahead of time and opening them up to a bribe or coercing them to act maliciously. The pool of potential candidates for the role of bookkeeper could contain both honest and malicious actors; it will be revealed later what would happen if the lottery happened to choose a malicious actor.

So, how would this lottery system work? Remember that the goal is to achieve a consensus mechanism that is fair, transparent, reliable, and efficient. This means that the rules need to be clear to everyone and the results easily verifiable. Summarizing Pritzker's illustration, the figure shows how the lottery would advance.



There are still a few missing parts of this system, the first being, how is the lottery administered without a central party? Who creates the lottery tickets and verifies that the winner actually has guessed the winning number?

Doing the Work—Introducing Miners

The short answer to the questions above is that the Bitcoin code creates the lottery system so no single entity needs to run it; it is part of the open-source code that anyone can view. Anyone who wants to participate in the lottery can do so as long as they follow the rules. If someone chooses to participate in this lottery and wants the chance to win a reward and add to the ledger, they are known as a "miner."

So, how do miners enter this lottery system? They instruct their CPUs to try to guess a randomly generated number that is not known to anyone ahead of time. It is literally a trial-and-error process, much like trying to guess the combination to a lock. Any miner who correctly guesses the combination gets the privilege of writing to the ledger and collecting a reward.

Why do miners perform this seemingly arbitrary task? As noted above, this is a way to make the lottery tickets costly in terms of computing power and electricity. If the lottery is open to anyone, but the tickets have no cost, then the system would break down because there would be no penalty or cost for malicious actors to overwhelm the lottery with tickets, or guesses, increasing their chance of winning and being able to write a fake transaction to the ledger.

> The physical cost of mining bitcoin is a feature of Bitcoin's system, not an undesired externality or byproduct.

As Pritzker notes in his book, electricity is always guaranteed to have a cost because it must come from somewhere (either produced by the miner or purchased from someone else), but it does not have to come from any central authority. In other words, this allows the lottery to be run in an efficient manner, open to anyone, without needing a central authority. Lottery tickets do not have to be created by a central lottery company; instead, anyone can create their own tickets by becoming a miner and choosing to expend computing power and electricity. The more power expended, the more tickets they are creating for themselves and the greater their chance of winning the privilege to write to the ledger and receive a reward.

So, how are the lottery winners verified and what keeps them from adding fraudulent transactions to the ledger?

The Asymmetric Nature of Proof-of-Work

The above example where miners perform a trial-and-error process of trying to guess a lock combination is an apt one because it demonstrates the task's asymmetric nature. This means that there is no possible way for the miners to get around the necessary work, while everyone else can easily verify that the work has been done. It may take trillions of guesses to open the lock (depending on how many digits are

needed to be found), but anyone can easily see that the correct lock combination was found because the lock is now open. A similar analogy could be a large jigsaw puzzle. It may take hours to correctly assemble a thousand-piece puzzle, but mere seconds for anyone else to verify that the puzzle has been assembled correctly because the picture is recognizable. **This is part of what allows Bitcoin to be a "trustless" system. Trust does not need to be put in a third party because anyone can verify for themselves that the puzzle has been solved and the work (electricity expended) has been done.**

This is how the lottery winners are verified as winners. A miner will announce that they have found the lock's combination and will then show everyone the solution. Everyone else can then take that proposed solution, enter it into the combination lock, and see that it indeed does open it. This is because each miner is working on the same lock, or problem. Once there is a verified winner, that miner gets to add to the ledger and the process starts all over with a new combination lock. This shows why this consensus mechanism is called **proof-of-work,** because providing the solution itself proves that work (trillions of guesses made by computers) has been performed. There is no other way around this; therefore, any valid solution is proof that work has been done.

The Incentive to Act Honestly—Game Theory

What if a miner legitimately wins the right to add to the ledger but then adds fraudulent transactions? Remembering that it is not a majority rules system, but this gives one entity the right to edit the ledger, couldn't this entity abuse this power? Technically, yes, they could try to add fraudulent transactions, but there is one final step. The rest of the participants must accept that batch of transactions as valid. In other words, the miner could go through all of the work of winning the lottery, only to see their final submission get rejected, and then they do not get the reward, nor do those transactions get added to the ledger. The next miner who won the lottery and proposed a batch of valid transactions would instead get the right to add those to the ledger and get the reward.

The physical cost of this system plays into how it aligns incentives for miners to act honestly. Because the energy must be spent before transactions can be added to the ledger, the miner is incentivized to act honestly to collect the reward for all their work that has already been done. In other words, there is already a sunk cost. It would be more profitable for the miner to use their computing power and electricity to collect the reward, rather than to try to propose adding invalid transactions that will then be rejected.

The Reward

Why do some participants voluntarily choose to enter this lottery if it costs computing resources and electricity? As previously mentioned, when miners successfully win and add a new batch of transactions to the ledger that are then approved by everyone else, they receive a reward in the form of bitcoin. This reward comes from two sources:

1. Transaction fees—all the transactions in the batch that are going to be added to the ledger have fees attached to them. The miners get to keep these fees as compensation.

2. The subsidy—the Bitcoin network has a built-in subsidy system where new bitcoin are created, or minted, and given to the miner who wins the lottery and enters valid transactions to the ledger.

At this point, it is only being illustrated that this system does not require the benevolence of miners to spend money to support the network. Later, this paper will examine some of the economics and implications of this incentive structure.

Security of Proof-of-Work (51% Attacks)

Because proof-of-work achieves a consensus, it can also be thought of as securing the network from errors (honest mistakes or multiple conflicting copies of the ledger) or deliberate attacks (someone trying to tamper with the true ledger so that they can double-spend bitcoin, reverse transactions, or block/censor transactions).

As previously noted, the proof-of-work system incentivizes malicious actors to act honestly, but if they still wanted to try to alter the ledger, then they would have to overwhelm the system with more computing power and, thus, electricity cost than all other participants combined. In other words, bad actors would have to overwhelm the lottery system with enough tickets to guarantee that they would be picked as the winner to then make a fraudulent entry to the ledger. To do this, they would have to control more than 50% of all the computing power currently dedicated to mining. This is known as a 51% attack and is a well-known attack vector for a decentralized network that uses proof-of-work as a consensus mechanism.

Summarizing Proof-of-Work

It seems abstract to think that computers consuming electricity to perform seemingly arbitrary number guessing can bring a decentralized network that nobody controls into an agreement or consensus, thereby also securing it from attacks. Yet this is exactly what proof-of-work does. To summarize:

1. The Bitcoin network is made up of many computers all running the same software code.

2. These computers all keep a copy of a ledger that shows every transaction ever made on the Bitcoin network. Transactions are messages sent out to these computers telling them to update the ledger (e.g. “move X bitcoin from this address to this address”). Therefore, the Bitcoin network can act as a payment network.

3. Because this network is decentralized with no one person or entity in control, it needs a consensus mechanism to come to an agreement on the ledger’s true state.

4. Proof-of-work acts as that consensus mechanism.

5. Proof-of-work creates a system that no one person or entity must run or control, but in which everyone can participate. This is done through requiring work or, more specifically, computing power and electricity usage. Anyone can do this work and use electricity—something that has a cost but does not have to come from a central authority.

6. Those who participate voluntarily choose to expend computing power and electricity for the chance to get to write or add new transactions to the ledger. For this service, miners receive a reward, which consists of a subsidy of newly created bitcoin plus all the transaction fees paid by the users sending transactions.

7. **Conclusion**—Proof-of-work is what allows the Bitcoin network to be decentralized and secure.

Governance and Proof-of-Work

Up until this point, this paper has only used the generic term "computers" for systems that run the Bitcoin software, which make up the network, and that some people run computers dedicated to entering this proof-of-work lottery. To understand the Bitcoin network's governance structure, what follows is a more detailed analysis of two different types of computers, or actors: nodes and miners.

Nodes

Nodes, also known as verifiers or verifying nodes, are computers running the core Bitcoin software. Anyone can download and start running the Bitcoin software, turning a computer into a node.[3] As part of the network infrastructure, nodes keep a copy of the ledger. When nodes receive new transaction messages, they update their copy of the ledger. Therefore, nodes have a full copy of all the transactions that have ever occurred on the Bitcoin network since its inception. As the operator of the core software and maintainer of the ledger, the nodes set and maintain the network's rules.

Miners

Miners are considered a special type of node that have voluntarily elected to enter the lottery and perform the previously mentioned number-guessing for the privilege of being the one to write to the ledger.

A walkthrough of a bitcoin transaction illustrates the different roles played by nodes versus miners:

1. A transaction is initiated by a user, who uses a piece of software that creates a digital message; the user sends that transaction message (e.g., "move X bitcoin from this address to this address") to the Bitcoin network.

2. Connected nodes receive and see this message, and check that it is valid (i.e., the message is not trying to create new bitcoin or double-spend previously spent bitcoin). If the message is valid, the

3 Note that the Bitcoin software is free and open source, allowing anyone to download it and run it without permission. The hardware requirements for running a node are very light and can be done with even a very low-powered, consumer-grade computer, making the electricity requirements to run a node very low.

node will pass it along to other nodes until everyone is aware of the message. At this point, the transaction is still pending and has not been confirmed (not yet written to the ledger). This smaller database of pending or unconfirmed transactions, which every node keeps, is known as the memory pool, or mempool. You can think of the mempool as a kind of waiting area for transactions wanting to be included in the next batch that gets written to the ledger. When a transaction is confirmed by being included in a block of transactions, it is removed from the mempool.

3. Miners select pending transactions from the mempool to be included in the next block of transactions to be written to the ledger.

4. Miners then compete, doing the guesswork to try to be the winner that gets to add that block and receive the reward.

5. The miner who correctly guesses the number then broadcasts that number and the block to the other nodes.

6. The nodes then verify that the broadcasted number is indeed the correct number, the work has been done, and relay the block to the other nodes. The block has now been confirmed and, therefore, the transactions included in that block are considered to have one confirmation.

7. If the miner did not follow the rules or tried to include fraudulent transactions in the batch, the other nodes would reject it, it would not be added to the ledger, and the miner would have wasted their time and cost of electricity. The process simply continues until another miner successfully solves the block, the proposed block is verified by other nodes and added to the ledger, and the miner receives the associated block rewards.

In the example above, transactions were first validated by the nodes before passing them to the miners. Miners then included them in the block they then worked to try to add to the ledger. The nodes then came into play again when they verified that the miner did the work (guessed the correct number) and accepted it to be added to the ledger by updating their copy.

Token Holders

For the actual holders of the bitcoin token (the asset), most of the actions mentioned above do not really apply in the same way as they do to nodes and miners.

You can see in the summary table on the next page where miners vs. nodes sit in the process and how each of their roles provides checks and balances to the system's governance. Notice that token holders have no governance role.

Bitcoin Network Participants Comparion Chart

	Nodes (aka validating nodes)	Miners	Tokens and Token Holders
Role/Purpose	Nodes make up the network; they: 1. Validate transactions. 2. Keep a record of transactions (the ledger). 3. Dictate and enforce the network's rules.	Write transactions to the ledger, therefore bringing consensus to a decentralized system; they: 1. Confirm transactions. 2. Secure the network. 3. Create, or mint, new coins.	Bitcoin are a provably scarce, natively digital asset that has characteristics of good money. Token holders subjectively ascribe value to bitcoin for these reasons and more, giving value to the network and incentivizing miners.
Governance Power	High—dictate and enforce the network rules.[4]	Low—must submit to and abide by the network rules or risk being ignored and not receiving any rewards.	None—the tokens themselves and holders of bitcoin have no governance power over the network.
Reward/Incentive	None directly, but anyone transacting or using the network has an incentive to run their own node to verify for themselves (not trust a third party) that the coins they are using are genuine and transactions are valid (i.e., make sure their transactions are safe and secure).	Monetary reward of newly created coins plus transaction fees.	Incentive for holding bitcoin is as an emerging money that can act as a store of value and/or medium of exchange. The reward to token holders may be increased purchasing power and asset price appreciation.
Cost	Relatively low—a low-end computer can run the node software ($100–$400).	High—capital-intensive equipment (>$10,000 per machine) plus electricity (high usage) and maintenance costs.	Cost of bitcoin is whatever buyers and sellers agree on; ongoing cost to hold is low to nil.
Risk and Complexity	Very little risk given low cost of equipment and no capital at stake. Complexity is relatively low if the node operator is only downloading and running the software, but can get more complex depending on use case or features needed.	High risk due to capital investment (that will depreciate and eventually become obsolete) and factors that can damage equipment (hardware failure, natural disasters, overheating, fire, theft, etc.). High complexity to set up, tune, and maintain machines in addition to constantly sourcing large amounts of cheap electricity.	Risk of loss may be due to operational error (lost keys) or loss of purchasing power (decline in price). Relatively low complexity compared with role of nodes and miners.

Why Can't Nodes Mount a 51% Attack?

Nodes are what make up the Bitcoin network by all running the same core Bitcoin software and, therefore, set the rules that everyone else (including the miners) must follow to participate in the network. Given this and, furthermore, that nodes are cheap, why can't someone mount a 51% attack by setting up many nodes?

Just like how the proof-of-work consensus is not majority rules voting, the nodes are also not majority rules. As a permissionless system, someone could indeed set up and control a majority of the nodes. However, to change the rules in their favor, they would need to propose a code change and the other (minority) nodes could then choose to accept and run the new code or stick to the old code. As the code is open source, anyone could see if a malicious actor was trying to propose new code to their benefit and not the benefit of the entire network and, therefore, would not adopt it.

4 For a real-world example of how nodes enforced network rules that were challenged, see *The Blocksize War: The battle over who controls Bitcoin's protocol rules* by Jonathan Bier.

Unlike some traditional majority rules political systems, the minority (good) nodes are not forced to adopt new code, even if a majority of other nodes do. The malicious actor would then be running their code alone while the rest of the network continued to use the original network, where consequently all the value and transactions would be occurring.

Implications of Proof-of-Work and Mining

Now that proof-of-work has been explained at a nontechnical level, the implications or secondary effects of this system and what they can mean for investors can be explored.

Linking the Digital World to the Physical

As alluded to above, the burning of real-world electricity is the core innovation of proof-of-work and one of the things that led to the breakthrough technology of the Bitcoin network.

Why the Link to the Physical World Can Give Bitcoin (the Token) Value

Furthermore, the implications are even greater for bitcoin the digital token. As written on in the past, bitcoin the token appears to be best understood as a monetary asset.[5] As such, it is important that it has a cost attached to it. While value is subjective, all else equal, it appears likely that people will attach a higher value to a monetary good that has an "unforgeable costliness" associated with it.

The term "unforgeable costliness" was most popularly mentioned in the context of money by computer scientist and Bit Gold creator Nick Szabo in a 2008 blog post.[6] The term is meant to describe things in the world that are costly (either due to taking a large amount of effort to create and/or find) and are difficult to spoof or fake the costliness. This is why the action of performing proof-of-work is akin to mining precious metals; there is a production cost associated with producing bitcoin.

Proof-of-work takes something that is completely virtual—a bitcoin token—and links it to the physical world. To paraphrase Knut Svanholm, bitcoin the token has no mass or matter and, therefore, is weightless.[ii] Bitcoin are merely entries in a ledger. Because of this, it makes them very easy to "move" across space and time, which is one of its most valuable properties as a monetary asset. However, this leads to the issue of, how do you make a weightless element "real"?

While bitcoin do not have any kind of mass or physical matter embodied in them, they do have energy "bound" or least linked to them. Because of the proof-of-work system, it is known that every bitcoin has come into existence through electricity expenditure. Proof-of-work is concurrently "proof-of-energy expended." In this way, you can think of each bitcoin token as a type of receipt that verifiably shows that electricity was expended to get that receipt. **Bitcoin has taken something in the real world (electricity usage) and verifiably transported it and permanently linked it into the digital world.**

5 See https://www.fidelitydigitalassets.com/research-and-insights/bitcoin-first-revisited

6 See http://unenumerated.blogspot.com/2008/08/

It should be emphasized how big of a technological leap this is for computer science and an increasingly virtual world. Everything seen on-screen is abstract and virtual; none of it exists in a physical form as everything perceived and interacted with on a computer is due to software, an abstract layer that tells the physical things (computer wires, chips, and circuit boards) to do certain things in terms of moving switches to "on" or "off" states. However, all this changed with Bitcoin, because there is now a completely virtual item (bitcoin) with proof that something physically exists (electricity).

Why the Link to the Physical World Can Help Keep Bitcoin (the Network) Decentralized

This link to the physical world also extends to help reinforce the decentralized nature of the Bitcoin network itself. The fact that mining requires large capital expenditures, sourcing of cheap power, electrical grid considerations, etc. means that the virtual Bitcoin network is forced to rely on these external factors and provides natural limits. Later, this will be contrasted with alternative systems that rely only on self-contained and purely virtual designated limits.

The key point is that to get a credibly decentralized network with no one person in charge, it may be more advantageous to have control or centralization limited by laws of physics. In other words, the reliance on physical electricity outsources some of the governance to the realm of physics and minimizes the role of human governance. Alternative systems that do not use energy must replace it with other forms of governance in order to limit centralization, which this paper will expand on later.

A Deeper Exploration of the Properties of Physical Power

It may still seem strange that the linchpin to getting a decentralized network to come to a consensus comes from burning electricity. However, a deeper understanding of electricity's physical characteristics can show how the use of it makes an excellent force for decentralization.

First, electricity is part of nature—it was discovered, not invented. As such, the laws of energy and electricity cannot be controlled by someone; therefore, one cannot be prevented from harnessing it and using it, a useful property to help run a permissionless and censorship-resistant network.

Second, electricity can come from many different sources. Therefore, this energy is distributed all over the world. Unlike oil-rich countries or rare-earth metals only found in certain parts of the world, electricity can come from using all kinds of material found everywhere: water, the sun, wind, even waves.[7] Electricity is also a perfectly homogenous commodity—there is no difference between electricity generated from wind versus natural gas, so one cannot exclude a type or grade of electricity. This is another useful feature to help run a decentralized network.

Third, there is no limit to the amount of electricity that can be produced in the world or used by the Bitcoin network. This limitless property is important because it means an attacker cannot monopolize a majority of the electricity market used in mining and gain an advantage that honest players cannot overcome. Instead,

7 See https://bitcoinmagazine.com/business/bitcoin-unlocks-ocean-energy

honest players can continue to buy, produce, and use electricity to defend the network against attackers. The attackers will then be faced with the pain of having to expend more resources or money to continue the attack. Since the honest participants will be willing to defend the network at nearly all costs, it is important that there is no artificial limit that the dishonest actors could reach before the honest actors.

Mining Promotes Fairness and Equality

The mining process is completely decentralized and open source. Furthermore, because of how mining works, there is no faking this energy expenditure. No one person or group can get around it and even Satoshi had to generate the first bitcoin by using electricity.[8]

This is not to say that early miners did not get more tokens for less amount of work (due to the reward schedule and something called the difficulty adjustment), but that is secondary to the main point[9] that even early miners had to follow the same rules and perform the same work. Token distribution is based on the amount of work done, regardless of the person or entity doing the work—the system does not discriminate.

Besides the token distribution being made fair through the process of mining, the proof-of-work consensus mechanism itself promotes fairness in adding transactions to the ledger. Again, because the mining process is distributed, decentralized, and open source, there is not one person or entity that can censor the transactions participants are seeking to add to the ledger. If some miners choose to prioritize adding some transactions over others or exclude some, other miners are incentivized to pick these up and add them in due time.

Mining Is Competitive, Not Cooperative

Earlier, it was mentioned that every miner works on the same problem, or combination lock, and the first to solve the combination lock receives the reward. At first glance, this may seem counterproductive as the other miners' work appears wasted. However, this is crucial to align the consensus mechanism's incentives.

The miners need to be competitive with one another to make sure that no party can collude or gain an unfair advantage. This is also necessary to keep the order of transactions, which is important to prevent double-spending. One miner cannot be working on a future combination lock and batch of transactions because they need the most recent one to work on the problem. This ensures that there is a chain of transactions that are verifiably linked together and are correctly ordered and time-stamped, which is a crucial element for a payment system.[10]

8 The discussion here is about Bitcoin specifically, which did not provide a pre-allocation of tokens. Proof-of-work does not necessarily mean that there are not pre-allocated tokens to certain founders or developers, as is the case with Ethereum, which started using proof-of-work but still allocated tokens to early investors.

9 Early miners receiving more bitcoin for less work could be thought of as an earned risk-premium given the incredibly nascent state of the mining industry at the time.

10 Mining pools are not an exception to this. When miners choose to pool their computing resources together, they are working *with* the mining pool, not *for* the mining pool. The process is still competitive as they compete against several other mining pools, and individuals within the same mining pool are still competing in a kind of zero-sum game where they only receive their share of rewards based on how much computing power they contribute to the pool.

Mining Can Be Anonymous, but Also Creates a Physical Footprint

Miners are just computers running the Bitcoin software and opting to enter the lottery or mining process. As such, they do not need to give up information identifying the operator or even the computer address or location (if set up in certain ways). This improves security and makes it more resistant to attacks. However, mining's physical nature does mean it needs capital equipment that is geographically located somewhere and connected to an electrical source. Large-scale mining operations are easily identified by their size and electrical consumption (for example, mining farms are similar to large data centers).

This raises some interesting observations when compared with other consensus mechanisms in terms of where attack vectors may form or the kind of centralization risks that may occur, as discussed later.

Common Questions and Myths on Mining

Mining Is NOT "Doing Complex Math Problems"

While miners perform calculations, these are not math problems, nor are they relatively complex. This characteristic gives the impression that miners are solving a difficult calculus problem and may lead to the misperception that someone else could come along and find a more elegant solution to outcompete other miners and, therefore, could use this more efficient solution to overpower the network.

An analogy would be the math student who is trying to solve for "x" in an algebra problem by "plugging and chugging" or using simple trial and error with different numbers until the solution is found, while another student goes through the algebraic steps to find the solution more quickly. Again, this is NOT how mining works.

Bitcoin mining is truly a plug-and-chug or trial-and-error system where the only way to find the solution is to guess different numbers. As such, there is no elegant way around this, and every miner faces the same difficulty level. However, miners can increase their odds of finding a solution by using a more powerful computer than a competitor or using a larger number of computers to increase their total power and, therefore, increase their guess rate.

If more computing power can increase the odds of finding a solution, would a supercomputer be able to mine bitcoin more efficiently or quickly? Not really, or only to a certain extent, and it would then be wasting its computing ability. Because the mining process is literally trial and error, specialized computer circuits and processors, known as ASICs (or application-specific integrated circuits), have been developed for the sole task of mining bitcoin more efficiently. Directing a supercomputer to do these relatively mundane guesses would work, but would not be the best use of the supercomputer's ability, which is to perform more complex calculations or actual complex math problems.

Since proof-of-work requires miners to be active to retain their percentage of power or tickets in the lottery system, miners are in a constant state of competition against each other. This means that they must constantly invest in the best mining equipment, access cheaper energy sources, and manage the

dynamic economics of energy supply and prices, as well as repair, maintain, and calibrate mining machines. Only the best, most efficient, and well-run miners are those that can maintain profitability, while any that become obsolete will be incentivized to shut down. Additionally, just because one entity might have a high percentage of hashing power currently does not mean it is guaranteed to stay that way.

Why Can't Mining Do "Useful" Work?

If it is essential for miners to do work, can they at least do something more useful than try to guess arbitrary numbers? For example, could they fold proteins, sequence genomes, or search for extraterrestrial life as part of the mining process while still participating in the proof-of-work process?

First, miners are doing something useful in that they are securing the Bitcoin network. Even if the task itself seems arbitrary, the result of the action is very useful and, as has been seen, is necessary to achieve consensus and provide security to a network that processes trillions of dollars of transactions per year.[iii]

Second, while it might be nice if the miners could somehow do another type of work that has a secondary benefit to others, it is hard to imagine how this would function in practice. Remember that the process should be open to everyone, with transparent and fair rules, and be able to be run without a central party in control. If miners were set to tasks of sequencing a genome, who would decide what genome to sequence next? How would the rules of winning be determined? Finally, remember that a key function of this mining process is that it is asymmetric; there is no possible way for the miners to get around the necessary work, while everyone else can easily verify that the work has been done. All these things would lead to the need for a third party to determine what is work and also judge when a miner has successfully completed the work.[11]

For those interested in the more technical aspect of this, miners actually perform a hash, or a one-way computing function where the miner adds an input (or a guess), trying to get a specific output (the winning number). These hashing functions were invented before Bitcoin and are well-known to computer scientists. Therefore, they make for an easy-to-verify, asymmetric system of proving work. For a more comprehensive technical overview, read our previous piece, "Bitcoin—Keeping Proof-of-Work Decentralized."

Miners Are NOT Processing Transactions

The idea that miners expend electricity to process transactions can lead to the confusion that it physically takes a large amount of electricity and computing power to calculate, or process, the bitcoin messages or transactions themselves. This is not the case. In fact, the amount of electricity needed to create, send, verify, and add transaction messages to the ledger is negligible in comparison to the proof-of-work electricity usage.

Another way to think about this is a network of low-powered nodes could handle all the processing of transactions with little electrical consumption if it did not need a consensus mechanism (i.e., it was controlled

11 A variation of this was attempted with Primecoin, a digital asset that sought to use finding prime numbers as its proof-of-work. Announced in 2013, it appears to not have made meaningful progress or adoption. See https://www.coindesk.com/markets/2013/07/10/new-currency-primecoin-searches-for-prime-numbers-as-proof-of-work/

by a single entity and was not decentralized). This is essentially what already exists with centrally controlled payment networks, like credit cards. Proof-of-work's high electrical consumption is what is needed to make the system decentralized and not dependent on third parties.

The secondary implication of this is that it is incorrect, or at least disingenuous, to calculate how much electricity is needed or used on a per-transaction basis. This would be like calculating how much electricity each email sent or received costs. In reality, the data servers that run email systems are running and using approximately the same amount of electricity whether just one or millions of emails are sent. Just as there is virtually zero marginal cost to sending an additional email, there is no additional electricity used to send more transactions over the Bitcoin network.

Proof-of-Stake

As previously noted, the Bitcoin network uses proof-of-work as its consensus mechanism. However, there are other consensus mechanisms that seek to fulfill the same purpose (maintain one ledger that everyone agrees on) but do so differently. One such alternative is proof-of-stake.[12]

While there are many different blockchains that use proof-of-stake as their consensus mechanism, and there are different variations of proof-of-stake, this paper will only focus on proof-of-stake as implemented by the Ethereum network and its native token, ether (ETH), the second-largest digital asset by market capitalization.[iv]

It is not the purpose of this paper to fully explain proof-of-stake in detail. However, because the purpose is to understand proof-of-work, reviewing and contrasting an alternative system can more clearly demonstrate the characteristics of proof-of-work and the different trade-offs between alternative systems that investors may face.

What Is Proof-of-Stake?

Like the above analogy of a lottery system in proof-of-work, only select winners get to add new transactions to the ledger in a proof-of-stake system. However, instead of performing trial-and-error guesses and competing to win the right to add to the ledger and receive a reward in a proof-of-work system, different validators are chosen at random[13] to write to the ledger in a proof-of-stake system.[v] Therefore, a proof-of-stake system has no miners, only validators. (These validators are not to be confused with validating nodes as in the Bitcoin network, as Ethereum also has separate nodes.)

Who gets to be a validator in proof-of-stake? Anyone who wants to, but they must first put up a stake of ETH and then join an activation queue that limits the rate of new validators coming online.[vi] So, instead of committing capital in the form of computing hardware and electricity, proof-of-stake requires capital in the form of ether tokens to be locked up and held as collateral with the risk of it being lost (more on this later).

12 Note that proof-of-work and proof-of-stake are being referred to as consensus mechanisms here, and while they are one large part of consensus mechanisms, there are other incentives and pieces of the protocol that make up the entire consensus mechanism.

13 Technically, these validators are selected pseudo-randomly, but for our purposes, the specific details are not important.

For the opportunity cost of locking up capital, validators on the Ethereum proof-of-stake system get newly minted ether tokens and "tips," or additional transaction fees paid by network users.

Similar to how a consensus mechanism was likened to a voting process earlier, proof-of-stake uses a more literal voting process. One validator is chosen every 12 seconds to propose a new block. This validator verifies that the transactions are valid, then sends the proposed batch of transactions around to all the other computers in the network with their vote, or attestation. There is also a committee of randomly chosen validators that then votes on whether the block that is being proposed is valid.

Security of Proof-of-Stake

As previously noted with proof-of-work, by providing a way to agree on the ledger's true state, the consensus mechanism is providing security to the network. Therefore, the end goal of security with proof-of-stake is the same, but achieved through a different set of incentives and mechanisms.

What Prevents Double-Spending in Proof-of-Stake?

In the original Bitcoin white paper, Satoshi noted one of proof-of-work's main value propositions was to prevent double-spending. This is done through the reality that if a bad actor wanted to change a batch of transactions, they would have to redo all trial-and-error guessing that was already done, spending a massive amount of computing power and electricity to do so. So, how does proof-of-stake prevent the changing of transactions or double-spending if there is no work to be done?

In general, proof-of-stake (at least as implemented by the Ethereum network) uses a system where attackers or malicious actors forfeit their staked capital (ETH) as a penalty. More specifically, it has regular checkpoints where validators vote for valid transactions to make them final. If in these checkpoints the batch of transactions gets votes that represent at least two-thirds of the value of the entire staked ETH, they are then finalized.

Therefore, if an attacker wanted to double-spend or otherwise change a transaction that had already been finalized, it would have to create an alternative record of transactions that are then voted on to be finalized and accepted as the true record. Given the voting rules above, this means that the attacker would need to own (or otherwise gain access to and control) two-thirds of the entire amount of staked ether in the network and be willing to destroy at least one-third of the total staked ether.[vii] At a price of around $2,000 per ether and 28.2 million ether currently staked as of November 2023, this would be equivalent to controlling $37.2 billion of ether, and burning a little over $18.6 billion worth of ether.

The specific details can get complex, but it is already apparent that there is generally an emphasis on voting and two-thirds majority rules with Ethereum's proof-of-stake system. Furthermore, because this voting is done by holders of ETH, it is similar to a system of shareholder voting and majority rules. The incentive structure is built on rewards of new tokens, similar to Bitcoin, but also on a system of penalties or risk of capital loss.

Summarizing Proof-of-Stake Attack Vectors

There is a wide array of attack vectors in Ethereum's proof-of-stake consensus mechanism—far too many and too complex to address here.[14] However, it is important to be aware of some of the generalities of these attack vectors. Additionally, there are some fundamental themes that are common with all these potential attacks:

- Given proof-of-stake's higher complexity and the larger amount of code it requires as implemented by Ethereum, there are a vast number of identified attack vectors both agreed on and demonstrated, but also some only theoretical or academic up to this point.
- There is ongoing and vigorous debate as to the likelihood of various attacks as well as the degree of damage that they could do.
- Because proof-of-stake's incentive system relies on putting up capital, most attacks require the attacker to have control of a certain percentage of the total ETH staked. While some researchers have demonstrated very nuanced and specific cases of using small amounts of capital, most attacks require control of at least one-third (33%) of the total staked ETH.
- Higher percentages of control of staked ETH create opportunity for more severe attacks that could give more control of the network to the attacker. For example, an attacker with control of one-third of total staked ETH could delay finality for a time, while two-thirds could control the network's past and future.
- The primary and often first line of defense is to slash, or reduce, the stake of an attacker and ultimately remove them from the network.
- However, many of the attack vectors, particularly the ones that are larger, more severe, or debilitating to the network, ultimately fall back on the social consensus layer to counteract or mitigate them—the coordination of Ethereum's people, users, and community. A real-world example of this was the DAO Hack, which relied on a community vote to roll back (fork) the code to eliminate the effects of an exploited bug in Ethereum's code.[viii]
- This greater reliance on the social layer to disarm attacks will be explored in more detail below when comparing proof-of-work with proof-of-stake.

Governance and Proof-of-Stake

In proof-of-work, the balance of power between miners and validating nodes was observed, where miners have no influence over the rules of the network and whose proposed blocks can be rejected by the validating nodes. Further, token holders of proof-of-work systems, such as Bitcoin, have absolutely no influence on the network rules or operation.

In Ethereum's proof-of-stake system, there are no miners, only validators who must pay the cost of locked or staked ether. There are still separate nodes in the Ethereum system that only run the software, but do not stake ether and do not propose blocks to be added to the ledger (nor do they get to write to the ledger). This

14 For a more detailed rundown of the various attacks, see https://ethereum.org/en/developers/docs/consensus-mechanisms/pos/attack-and-defense/

is similar to the Bitcoin nodes in that Ethereum nodes uphold the rules and also propagate valid blocks of transactions, while rejecting invalid ones. While Ethereum nodes do require a relatively more powerful computer with higher minimum specifications than Bitcoin nodes, they are still within the realm of higher-end consumer computers that can be built or purchased off the shelf.[ix]

However, while this structure is technically the same as with Bitcoin, in practice, it has been observed there is not nearly as much of an emphasis on running your own node in the Ethereum community as there is in the Bitcoin community and, in reality, it appears most nodes are run by developers, enterprises, and stakers, rather than individual holders and transactors in Ethereum.[x]

In summary, the Ethereum network has three major participants: nodes (like Bitcoin nodes), validating nodes (instead of miners), and token holders. **However, the roles of the validating node and the token holder now overlap.** The next section will explore some of the implications of this.

Finally, it should be emphasized that, while validating nodes practice a more literal voting process in their validation, this is not the same as voting for changes in the protocol or improvements. Validating nodes and, therefore, stakers can only vote on which blocks and transactions are added to the ledger, with more tokens staked exerting more control. But changes to the protocol are done through Ethereum Improvement Proposals (EIPs) and social consensus through communication with the developers, foundation, and community, which must ultimately be adopted by the nodes running the software.

In a strict technical sense, more tokens held or staked does not equal more votes when deciding on changes to the protocol or Ethereum roadmap. However, some critics may point out that, in practice, the Ethereum Foundation and other large token holders carry a lot of sway in these matters and their changes to the protocol have been more forceful. For example, Ethereum's use of difficulty bombs[15] versus Bitcoin emphasizing things like soft forks.[16]

Comparing Proof-of-Work with Proof-of-Stake

The main focus of this paper remains understanding proof-of-work as part of the Bitcoin network's consensus mechanism. While understanding proof-of-stake is secondary, comparing and contrasting the two systems can help demonstrate the trade-offs made and the advantages and disadvantages of each.

Since there are numerous ways to compare the two, the comparison is limited to a number of specific vectors or lenses of evaluation that appear to be most important.

15 "Difficulty bombs" refers to a sudden increase in mining difficulty to discourage miners from opting to stay with the proof-of-work mechanism after it transitioned to proof-of-stake.

16 Soft forks are changes to a project's source code, which are backwards compatible, meaning that it is not necessary that all nodes adopt the upgrade. Soft forks preserve consensus and flexibility in a distributed system because they allow some nodes to adopt upgrades and make use of new innovations while other nodes who do not wish to upgrade are not forced out of the network or put at risk.

Complexity

It is widely agreed and observable that proof-of-work is simpler than proof-of-stake from a technical perspective. Furthermore, even Ethereum developers have noted that proof-of-stake is more complex than proof-of-work, and that it is harder to introduce bugs or unintended effects into simpler protocols.[xi] This paper is not intended to explore all the technical reasons for this, but even from the previous sections, it is clear that proof-of-stake introduces potential problems not found in proof-of-work that must be addressed through more rules and governance, such as slashing or reducing stakes and ejection from the network.

Proof-of-stake was supposed to be the mechanism securing Ethereum from the start, according to its original white paper published in 2013.[xii] However, as Ethereum Founder Vitalik Buterin noted in 2014, developing such a system was "so nontrivial that some even consider it impossible."[xiii] Ethereum launched with a proof-of-work model instead and began developing a proof-of-stake blueprint. This technical complexity is also why Ethereum's switch to proof-of-stake took years longer than anticipated.

Of course, complexity in and of itself is not necessarily a bad thing. Phones have become significantly more complex over the past couple of decades. What was originally just capable of simple phone calls—not even leaving voicemails—has evolved into a handheld phone, computer, GPS system, game console, music player, and more.

Investors should not necessarily shy away from complex investments and it should be acknowledged that the Ethereum network has successfully transitioned to proof-of-stake with no bugs, interruptions, or adverse events during or since the transition. However, investors should be aware of the degree of relative complexity and know that increased complexity may increase risk, higher probabilities of failure and bugs, or broader potential attack surfaces.

Decentralizing vs. Centralizing Factors

Starting with the assumption that decentralization is the goal and a preferred attribute of both proof-of-work Bitcoin and proof-of-stake Ethereum, one must look at the general term "decentralization" through many different perspectives.

Supply Chain and Economies of Scale

Critics of proof-of-work will point out that the supply chain behind the mining machines needed for competitive proof-of-work is fairly centralized with only a handful of ASIC[17] manufacturers. Large mining operations can outcompete others by securing favorable energy or computer chip contracts, or perhaps even manufacture their own proprietary chips (i.e., vertically integrate).

Furthermore, at an even more basic level, the degree to which bitcoin mining has advanced means only relatively deep-pocketed individuals or, more realistically, companies, can mine for bitcoin as machines cost

17 Application-Specific Integrated Circuits (ASICs) are purpose-built machines that are designed to only perform the task of mining bitcoin or other proof-of-work protocols.

tens of thousands of dollars each and require thousands of watts of energy. This favors larger players that can operate at greater economies of scale with large mining farms and the ability to buy machines and electricity in bulk.

In contrast, almost anyone can stake ether (especially with the use of ether staking pools) with no specialized hardware or favorable electricity prices needed. Each person who puts in a unit of capital in the form of ether is rewarded the same as the person who puts in multiple times more capital.

The supply chain is likely one potential risk to proof-of-work, especially in terms of Bitcoin, and is something to continue to monitor. Further, it is not feasible for most individuals to mine bitcoin anymore (at least for a reasonable profit or expected amount of bitcoin). However, advocates for proof-of-work would point out that, while there are indeed large bitcoin mining companies, they are still only a fraction of the network's entire hash rate or total computing power,[xiv] making the security concern of too much concentration low. Additionally, because miners have no power over the network rules and protocol, and they are involved in a competitive process, the only concern is that miners or their supply chains do not take control of more than 50% of the entire hash rate or computing power.

Critics of proof-of-stake will point out that, while individuals can stake ether, it is currently a relatively complex process requiring some technical expertise. Therefore, most individuals will turn to staking services or third-party services provided by centralized exchanges or custodians. Financial history and the current financial landscape show that custodians and exchanges are relatively few in number with only a handful custodying the majority of financial wealth. Therefore, it is reasonable to think that there may be similar concentrations established in staked ether and an increasing tendency toward an oligopoly of stakers.

Jurisdiction and Geographic Concentration Risk

Both proof-of-work and proof-of-stake face jurisdictional and/or geographic concentration risks, but in different ways. Proof-of-work obviously has a physical footprint, particularly with larger mining operations that have entire farms, similar to large data centers. These can be roughly 100,000 square feet or more in size and need access to large power sources. These facilities require large amounts of capital and cannot quickly and easily pick up and move because physical equipment needs to be packaged, moved, set up, and plugged in at a new facility with more favorable electricity prices.

Proof-of-stake as a completely virtual endeavor does not have these same physical issues. However, custodians, exchanges, or other staking service providers do face jurisdictional risk and could be subject to unfavorable laws, regulations, or entire capture/nationalization. While this may not be a likely scenario because the roles of validators (stakers) and token holders can overlap, there is relatively more loss in the event of an attack and/or loss of control for large custodians or staking services compared with proof-of-work.

For example, if an entire custodian of ether that was staking ether was hacked, co-opted, or nationalized, the network would lose the security of that staked capital and those who owned that staked ether would lose their assets. The same risk can apply with proof-of-work, but they would have to be separate attacks

(i.e., losing control of a large mining farm would not necessarily mean the bitcoin holders would lose access to their coins).

The Rich Get Richer?

In a proof-of-stake system, once someone stakes capital to the network, they continually receive new tokens and transaction fees as compensation as a pro rata share of the total rewards, depending on how much they stake out of the total amount being staked. This has a few implications, the first being that anyone who holds tokens, such as ether, but does not stake them, will continually see their share of the total ether supply being diluted, while the stakers do not.

This obviously creates a strong incentive for anyone who wants to hold ether for a longer time period to stake them. This in itself is not a normative statement, but a more opinionated perspective is that it could be good because it incentivizes holders to help secure the network. On the other hand, it could have unintended consequences of pushing more people at the margin into staking services and pools, creating large staking oligopolies, as mentioned earlier.

The second implication is the idea that the rich get richer, or that anyone who amassed or was allotted a large amount of ether tokens early and is now staking them, will never see their overall percentage of staked ether decline. In other words, nothing can dilute their power or influence over their votes in the validating process, which decides which blocks and, therefore, which transactions are allowed to be added to the ledger.

Proof-of-work, on the other hand, requires miners to always be active to retain their percentage of hash power, or lottery tickets, that ultimately determine who gets the privilege of updating the state of the ledger. Miners must compete against other miners and must continually invest in new equipment, source or secure favorable power sources, and deal with power demand or price changes as well as repair, maintain, and tune new mining machines.

In practice, this means that the best, most efficient, and well-run miners are those who are being rewarded and survive, while any who become unprofitable could be driven out of the business. This also means that just because someone or some company might currently have a high percentage of hashing power does not mean that it is likely to stay that way. Contrast this with the passive approach of staking, where large stakers can remain large stakers indefinitely.

Governance at the Social Layer

Ultimately, both proof-of-work and proof-of-stake are only one part of the consensus mechanism and the entire blockchain ecosystem that they support. One other major component of consensus for both the Bitcoin and the Ethereum networks is the social layer, or the coordination and consensus between people and the community.

For example, all bitcoin token holders agree to use the current Bitcoin network, its code, and its proof-of-work consensus mechanism that employs mining. Governance has also transpired at the social layer with

the Bitcoin network as it has gone through the adoption (and rejection) of different Bitcoin Improvement Proposals (BIPs) or code changes and upgrades. These changes came about through communication with the community, developers, token holders, and miners with lots of debate and signaling of preference. Ethereum also has a similar, and possibly even stronger, social consensus layer.

What does this have to do with proof-of-work versus proof-of-stake? While it is beyond the scope of this paper, when it comes to attacks against the network's consensus mechanism, proof-of-stake as employed by Ethereum relies much more heavily on using the social layer as a means of thwarting a potential attack. This includes having the community band together to decide how to effectively deal with an attack, such as in what form (slashing, network ejection, freezing of funds, destroying entire stake) and to what degree of severity.

Proof-of-work with Bitcoin can also rely on the social layer (for example, honest miners could try to band together during an attack to bring on more computing power), but the available options are somewhat fewer than proof-of-stake because the proof-of-work consensus mechanism and its rules are more objective.

This is not to say that relying more or less on the social layer is better or worse, but to point out the differences with each system. However, there are trade-offs between the two and it is important to use the system that is best suited to the task or objectives at hand.

Physical vs. Virtual Anchors

A consensus mechanism's goal is to have a trustless and decentralized system and, therefore, to find a single trustworthy person or entity that gets the privilege of writing to the ledger. With proof-of-work, trust is given to the user who provably spent much effort and energy. In proof-of-stake, the user is considered trustworthy because they have a significant stake in the system.

Both proof-of-work and proof-of-stake hinge on requiring a cost to work properly and both ultimately require a cost of capital: an opportunity cost of using capital to help secure a network when that capital could be used for other purposes. With proof-of-work, the cost of capital comes from the money invested and tied up in mining machines and the ongoing money used to buy electricity to power them.

With proof-of-stake, the cost of capital is more direct in a sense as it is locking up the capital itself, such as ether, rather than relying on more steps of buying machines and electricity. But in another sense, it is more indirect or fuzzy as the capital required is native to the network it itself is trying to secure. In other words, the Ethereum network is self-contained and relies on its own network and systems to uphold the security and consensus, while the Bitcoin network has a link to the physical world.

Another way of looking at this is akin to the previous discussion on social layers being used to achieve consensus. While both systems use the social layer, the Bitcoin network has an additional layer, a physical layer to help achieve security and consensus, while the Ethereum network has removed this physical layer.[18]

18 Note that Ethereum used to have this physical layer as it was first built using proof-of-work and, as of September 2022, successfully switched to proof-of-stake with the event known as "The Merge."

Ethereum's prominent co-creator, Vitalik Buterin, summarized this difference:

> *"Proof-of-work is based on the laws of physics, and so you sorta have to work with the world as it is...you have to work with electricity as it is, hardware as it is, what computers are, as it is. Whereas because proof-of-stake is virtualized in this way, it's basically letting us create a simulated universe that has its own laws of physics, and that just gives us as protocol developers a lot more freedom to optimize the system around actually having all of the different security properties we want... If we want the system to have a particular security guarantee and then often there is a way to modify the system to also achieve it..."*[xv]

Without assuming or inferring anything from Buterin in this quote, it is a good overview of the trade-offs made between these two systems and reiterates the earlier point that proof-of-work's link to the physical world is a "feature, not a bug" when employed to secure something where its holders or those who value it want it to be difficult to change the underlying consensus mechanism.

Weak Subjectivity or the Role of Trust

When referring to these networks, the term "subjectivity" refers to relying on social information to agree on the ledger's current state. Conversely, "objectivity" would refer to networks that do not require social information and where the current state is completely known by using the coded rules and software's logic. An in-between term, "weak subjectivity," refers to when a network needs to get an initial start from social information, but then can proceed more objectively.

These terms are used more with proof-of-stake as subjectivity is inherent to proof-of-stake because the ledger's correct state is determined by counting historical votes and exposes the system to several attack vectors.[xvi] Without getting too specific, this is why proof-of-stake networks' security requires nodes to be connected and online as much as possible, whereas if proof-of-work nodes are offline for any period of time, it is easy for them to determine what the ledger's correct state is when they return online without risk of being fooled or attacked.

The other big difference is that with Bitcoin and proof-of-work, every node can download the entire Bitcoin ledger and verify for themselves the history of every bitcoin token ever created and every transaction completed, all the way back to the original "genesis block," or the very first bitcoin block, mined in early 2009.

Ethereum using proof-of-stake does not do this, but rather implements "weak subjectivity checkpoints," or certain states and points in time, which everyone agrees are the source of truth. The risk, of course, is that participants must then trust that these checkpoints are indeed valid and that they are getting these checkpoints from other honest participants. The Ethereum network seeks to minimize this risk by checking the checkpoint against other independent sources and having multiple independent teams that are building equivalent software in different programming languages, and assuming they all have a vested interest in building an honest ledger.

The Bitcoin network does not need these checkpoints, but critics do point out that there are other pieces of information that the Bitcoin network needs that must be acquired from other participants (i.e., other socially derived pieces of information). This includes from where someone downloads the core Bitcoin software as well as the supply chain of their hardware and other software, such as the operating system. While this is technically true, it is not of the same magnitude of risk as proof-of-stake weak subjectivity and is also the same risk that every other blockchain faces.

This is not a proposal of probabilities on how high or low this weak subjectivity risk is, but it is important to note as a key difference in terms of what degree of trustlessness is needed for each mechanism.

Positive vs. Negative Incentives

As seen with proof-of-work, honest miner behavior is incentivized by rewards in bitcoin, while malicious activity is disincentivized by miners wasting energy. Proof-of-stake also rewards its validators with ether tokens, but the incentives to act honestly come more from avoiding the negative aspect of slashing or burning staked ether. This is not to say that a system of actively slashing or reducing a validator's stake cannot incentivize honest behavior, but it is nevertheless something to be aware of in case situations arise where validators may be incentivized to side with the majority voters for fear of losing their stake.

Comparison Chart

Proof-of-Work		Proof-of-Stake
Mining ability based on computational power	**Mining/Validating**	Validating ability based on stake in the network
Miners receive block subsidy (new bitcoin) and transaction fees	**Reward Distribution**	Validators collect attestation (voting) rewards, transaction fees
Miners compete to guess a randomly generated number using computer processing power	**How a Winner Is Chosen**	Randomly chosen by the algorithm
Cost comes in the form of dedicating hardware, electricity, ongoing maintenance	**Capital Cost**	Capital is tied up directly by staking tokens
Requires specialized equipment to optimize computing power	**Hardware Requirements**	Standard server-grade equipment
High energy demands	**Efficiency**	Negligible energy demands
Slower transaction speeds (10–60 minutes to transaction finality)	**Speed**	Faster transaction speeds (13 minutes to transaction finality); also, switch to proof-of-stake is first step to eventually increase speed
Historically proven, longest track record of any digital asset	**Track Record**	Ethereum is now largest proof-of-stake network, but track record is short thus far (since Sept. 2022)
Token holders have no power over which transactions are added or censored	**Governance of Token Ownership**	Larger token holders that stake have more influence over which transactions are added or censored
Initial investment in buying the hardware	**Startup Capital Required**	Initial investment in buying the stake
Attackers need 51% of total network computing power	**Malicious Attacks**	Varies from needing control of at least 33% of total stake to 66% for major attacks, plus willingness to burn significant portion of stake
Token holders have no governance power, miners have power over transaction inclusion, nodes have power over protocol rules	**Governance Structure**	Nodes have power over protocol rules, validators have power over transaction inclusion, token holders can be validators through staking
Lower	**Complexity Level**	Higher
Lower—can audit entire ledger back through its full history	**Subjectivity Level**	Higher—requires weak subjectivity or more trust in other honest nodes or participants
Yes—relies on laws of physics in terms of power, electricity, etc.	**Physical Governance Layer**	No—completely virtual
Lower	**Social Governance Layer**	Higher

Conclusion

This paper shows how proof-of-work operates as a consensus mechanism, the implications of such a system, how the competing proof-of-stake operates, and a comparison of proof-of-work vs. proof-of-stake to try to better understand each mode of reaching consensus more fully.

It is not the object of this paper to declare proof-of-work as superior to any other type of consensus mechanism, just as it would be foolish to declare a four-door sedan as the best type of vehicle. The correct approach is to first examine what the objectives are and then find the best tool or mechanism to optimize for those objectives.

In a previous report, "Bitcoin First," we proposed that, "Bitcoin is best understood as a monetary good, and one of the primary investment theses for bitcoin is as the store of value asset in an increasingly digital world. Bitcoin is fundamentally different from any other digital asset. No other digital asset is likely to improve upon bitcoin as a monetary good because bitcoin is the most (relative to other digital assets) secure, decentralized, sound digital money and any 'improvement' will necessarily face trade-offs."

With a deeper understanding of the proof-of-work that underpins bitcoin, it becomes clear how this unique consensus mechanism contributes to making bitcoin the most secure, decentralized, and sound digital monetary good. Most notably, it appears likely that proof-of-work could provide the most fair, transparent, and decentralized consensus mechanism.

Bitcoin's link to the physical world and use of real-world resources is one of its primary features and competitive advantages that lends value to the token and network, while also enhancing and solidifying its role as a digital commodity and emerging monetary good. In short, proof-of-work's properties produce and strengthen the core value proposition of Bitcoin, creating a beautiful alignment of incentives and self-reinforcing characteristics.

This is not to say that proof-of-stake as a consensus mechanism cannot succeed, and it should be noted that the second-largest digital asset network, Ethereum, not only successfully transitioned from proof-of-work to proof-of-stake, but did so with no interruptions or bugs in the process or thereafter.

However, as explored above, the proof-of-stake consensus mechanism makes different engineering trade-offs, such as higher complexity, more reliance on the social layer and community for security, the introduction of weak subjectivity, and potentially centralizing forces.

Proof-of-stake removes the reliance on the physical layer as a governance and incentive mechanism, which may be more appropriate for assets that have and want to continue changing, whereas the physical demand and limits of proof-of-work appears to be perfectly suited for an asset whose main value proposition is to be relatively non-changing and ossified, one that can be counted on to be the same years, decades, or perhaps even centuries from now.

i Yan Pritzker, *Inventing Bitcoin: The Technology Behind The First Truly Scarce and Decentralized Money Explained,* (Independently Published, 2019).

ii Knut Svanholm, *Bitcoin: Everything divided by 21 million*, (Konsensus Network, 2022).

iii Coin Metrics.

iv Coin Metrics.

v See https://ethereum.org/en/developers/docs/consensus-mechanisms/pos/

vi See https://ethereum.org/en/developers/docs/consensus-mechanisms/pos/#validators

vii See https://ethereum.org/en/developers/docs/consensus-mechanisms/pos/gasper/ and V. Buterin, D. Reijsbergen, S. Leonardos and G. Piliouras, "Incentives in Ethereum's Hybrid Casper Protocol," 2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC), Seoul, Korea (South), 2019, pp. 236–244, https://doi.org/10.1109/BLOC.2019.8751241

viii See https://ethereum.org/en/developers/docs/consensus-mechanisms/pos/attack-and-defense/

ix See https://ethereum.org/en/run-a-node/

x See https://bitnodes.io/ and https://www.ethernodes.org/

xi See https://ethereum.org/en/developers/docs/consensus-mechanisms/pos/pos-vs-pow/

xii See https://ethereum.org/en/whitepaper/

xiii See Buterin's full blog post from 2014 https://blog.ethereum.org/2014/10/03/slasher-ghost-developments-proof-stake

xiv See https://mempool.space/graphs/mining/pools

xv Vitalik Buterin, *Ethereum Mainnet Merge Viewing Party*, recorded live September 14, 2021, timestamp 2:31:20, https://youtu.be/Nx-jYgI0QVI?t=9071

xvi See https://ethereum.org/en/developers/docs/consensus-mechanisms/pos/weak-subjectivity/

1104856.1.0