RESEARCH REPORT ✷ JUNE 2024

Ethereum's Correlation Risks: Poorly Understood, but Always Present

A holistic view of today's diversity and correlation risks on Ethereum, plus recent proposals and upcoming changes that may shift network dynamics

Obol × Liquid Collective

Contents

WORDS
Max Sherwoo — Content & Communications Manager, Obol Labs
Melissa Nelson — Head of Content, Alluvial

DESIGN
Mark Forscher — CMO, Alluvial

THANK YOU TO OUR CONTRIBUTORS
Editor, editor

Summary

BACKGROUND — A healthy diversity of nodes and validators is important for a blockchain network to create a globally-accessible network of peers, with constant uptime, correctness, and finality. What makes Ethereum unique is that slashing penalties increase, in the event that an entire portion of the network is acting harmfully—if there is a "correlation."

PROBLEM — For any staker delegating to a single, large, undiversified operator, correlated slashing represents a very rare, but very real, risk, with large implications. Active proposals indicate that the correlation risks on Ethereum are unlikely to fade in importance. Further correlation disincentives beyond correlated slashing may be added, increasing the importance of staking with decentralization across operators and infrastructure.

SOLUTION — Examining a holistic cross-section of today's correlation risks on Ethereum, along with how they may shift tomorrow, can support informed decision-making within a diverse staking ecosystem. Key areas for evaluating correlation risks in staking include client diversity, and the supermajority client concern, node operator diversity, and cloud and regional diversity. Key correlated risk mitigations include staking across a diverse set of node operators, diversifying validators across multiple clients, clouds, and regions, and leveraging distributed validator technology (DVT) strategies.

1.

Why do Ethereum's Correlation Risks Matter?

Blockchains Need Diversity

Not all blockchains are created equal. Usually, they start with a very small set of validators, with the tokens held by a small group, before slowly diversifying and distributing themselves more broadly. The goal is to create a global, decentralized ecosystem that is resilient against any adversity. Arguably, that work is never finished.

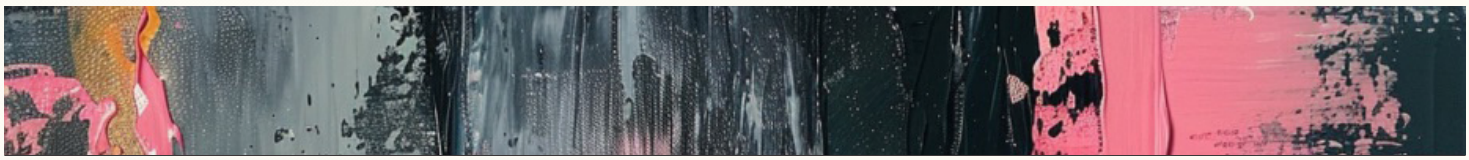
A healthy diversity of nodes and validators is important to create a globally-accessible network of peers, with constant uptime, correctness, and finality. The network should continue to operate no matter the circumstance, if it is to be adopted by applications and used for financial activities at scale. Economically, a diversity of stake is also important, as it ensures the network cannot be attacked or taken advantage of by a single entity or small group. While generally, an increasing amount of stake in a network increases the cost of attack, not all stake adds the same "usefulness" to the network, if it doesn't contribute to the network's diversity.

Many participants think about staking in terms of the reward rate received for adding credibility to a validator's work, but it's more than that: staking is playing an active part in securing a global decentralized network. Behind every staker is a validator node performing critical duties. If someone is willing to invest in supporting Ethererum's operational infrastructure, for example, they should be aligned with the ability of Ethereum's infrastructure to survive into the next 100 years and become the next iteration of the web, and contribute as best they can.

Today, Ethereum has thousands of nodes spread across the world and broadly distributed token ownership, making it adequately decentralized to withstand single points of failures or attacks. However, very real risks remain, which stakers should understand.

Correlation is a crucial area to develop this understanding for the risk-adjusted evaluation of staking technologies. Some of today's hot protocol research topics include diversity, the target staking rate percentage, and enshrining methods to incentivize more decentralized node operation. Meanwhile, the community calls for stakers to adjust their strategies and creates a range of Ethereum Improvement Proposals (EIPs) focused on addressing different angles of correlation in staking.

Examining a holistic cross-section of today's correlation risks on Ethereum, along with how they may shift tomorrow, can support informed decision-making within a diverse staking ecosystem.



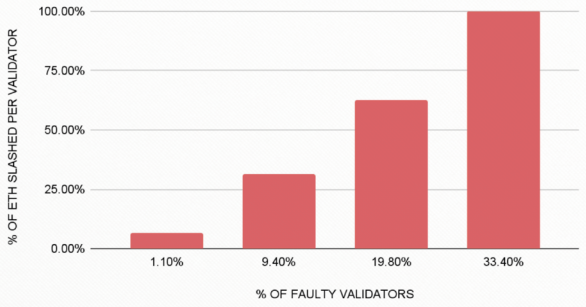
Are Today's Correlation Penalties Working for Ethereum?

Staking is a system of carrots, and in Ethereum's case, sticks. Staking rewards incentivize node operators to keep their validators operating effectively, while offline penalties do the inverse. Beyond that, Ethereum has a powerful slashing mechanism that disincentivizes poor validator behavior, like double-signing, which is harmful to the network. What makes Ethereum unique is that these penalties increase, in the event that an entire portion of the network is acting harmfully—if there is a "correlation."

What happens when a fatal flaw with a client causes correlated slashing? Instead of the usual slashing penalty of 1 ETH of the 32 ETH in a validator, a validator will lose 25 ETH if it is being slashed alongside 25% of the stake on the network. Put simply, the more validators that are slashed at once, the worse the penalty per validator becomes.

Vitalik puts it best when he explains that "the theory is that if you are a single large actor, any mistakes that you make would be more likely to be replicated across all 'identities' that you control." This penalty supports the resilience of the network by increasing risks for those who represent a potential "single point" of failure, effectively economically disincentivizing a malicious actor from taking over control of the network.

Correlated Slashing Scenarios



Ethereum severely penalizes correlated slashing events:

- If 1.1% of the network is slashed: 0.1% of funds are lost
- If 33.4% of the network is slashed: 100% of funds are lost

SOURCE: LIQUID COLLECTIVE

What many stakers don't consider is that this validator activity need not be malicious. Given Vitalik's point—that one operator running many nodes (without sufficient distribution) is likely to execute the same staking practices across all of those nodes—excessive downtime, bugs in client software, geographical outages, and a plethora of other non-malicious causes could lead to 1.10% or more of the network experiencing a correlation event (the minimum percentage of the network's ETH slashed to kick in the correlation penalty).

Overall, as an economic disincentive, Ethereum's correlation penalty is intended to provide resilience against any one entity controlling so much of the network that their mistake or malicious behavior could cause significant harm. This shifts the impetus of ensuring that the Ethereum network is able to reach finality and stay active onto its operators: the risk of loss becomes too great to take diversity lightly at scale.

COMMUNITY PUSH TO REDUCE CORRELATION

The threat of correlated slashing has encouraged node operators to be wary of running consensus or execution clients which are also used by the majority of the network. In January 2024, a Nethermind client bug brought client diversity concerns back to the forefront after many in the community noted that a similar bug in Geth, which operated on 84% percent of the network at the time, would cause Ethereum to stop reaching finality, causing other waterfall effects.

The resulting ecosystem-wide call to action for the staking industry to diversify has led to the more evenly-distributed adoption and use of five consensus clients and four execution clients. But with large percentages still using individual clients, risks remain.

The risk of correlated slashing is rare, but it does exist, and for good reason. But is it effective in its goal to encourage decentralization? Or may protocol-level upgrades be required to more effectively incentivize stakers to reduce correlation across all of its vectors? An economic disincentive for a long-tail risk that isn't respected or acted-upon by the majority actually may just raise risks in the system, as it represents the non-zero potential of billions of dollars in staked assets being burned, should sufficient mitigations not be in place.

Vitalik made an interesting proposal in late March 2024, suggesting that the correlated slashing mechanism was not achieving its objective, and that further correlated penalties might be necessary—for example, around missed attestations.

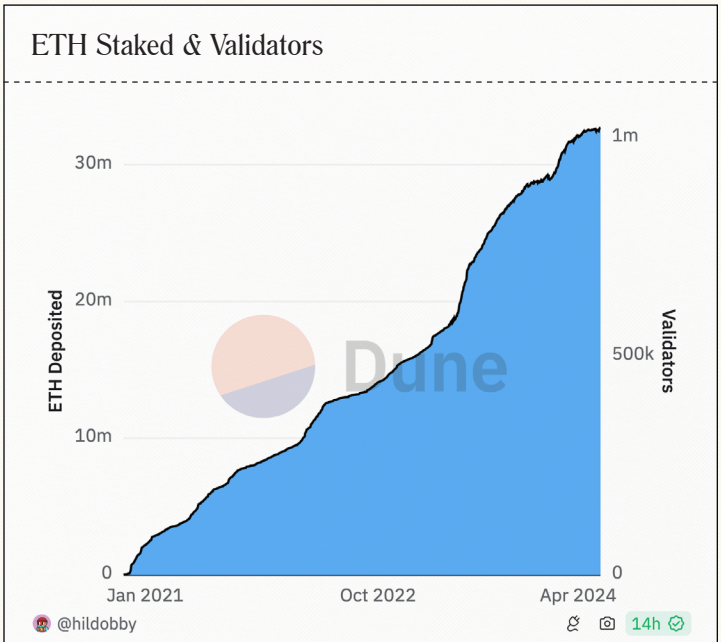
While a meaningful correlated slashing incident represents more of an "existential risk" scenario—one that could potentially break down participant trust and value in ETH on a macro scale—missing attestations represents a much more common, menial form of "poor" behavior. This penalty for missed attestation correlation would further increase risks around staking with large node operators whose infrastructure is not sufficiently decentralized, by negatively impacting their staking reward rate (SRR) over time, though their poor practices may never rise to the level of causing a full correlated slashing incident.

We'll explore Vitalik's proposal, and other proposals surrounding correlation, in the "what's next" section at the end of this report. But, based on these recent conversations, it appears clear that the correlation risks on Ethereum are unlikely to fade in importance. Further correlation disincentives beyond correlated slashing may be added, increasing the importance of staking with decentralization across operators and infrastructure.

THE STAKES KEEP INCREASING

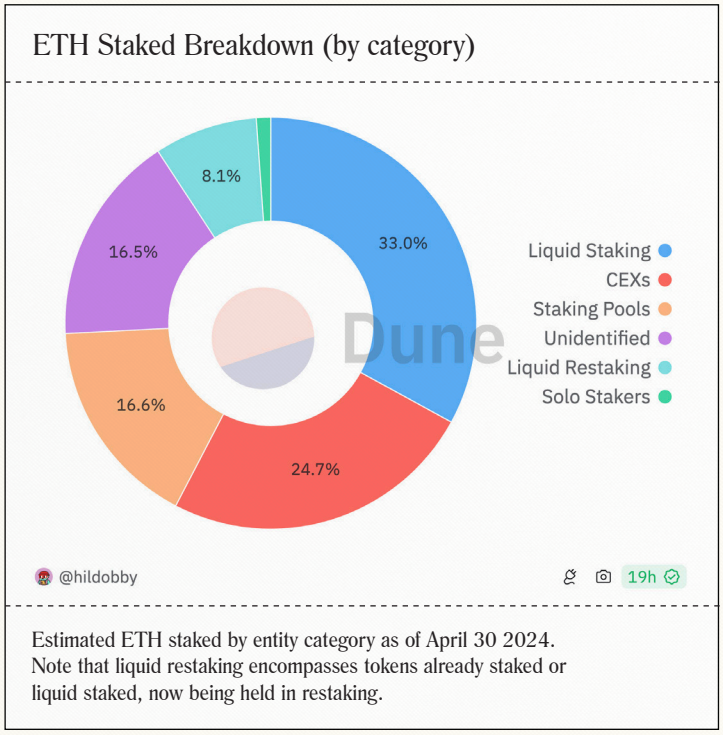
Staking has grown substantially in recent years: in 2023, the amount of ETH staked grew from 15.8m ETH ($18.97b at the time) to 28.9m ETH ($63.5b today), over 3x in dollar terms. As the Ethereum protocol has evolved, integrating new features like staking withdrawals, ETH holders have become more comfortable staking their ETH, a trend which continues.

Estimates for what percentage of staked ETH is staked by solo operators vary, but it's clear that they represent a minority. According to estimates from onchain analyst Hildobby, just 1.08% of all staked ETH is staked by solo operators today, while Rated estimates 6.5% of validators are run by solo stakers.



ETH staked over time.

SOURCE: HILDOBBY - ETHEREUM STAKING - DUNE



Estimated ETH staked by entity category as of April 30 2024. Note that liquid restaking encompasses tokens already staked or liquid staked, now being held in restaking.

SOURCE: HILDOBBY ETHEREUM STAKING DUNE

Solo stakers, also called solo operators or home stakers, is a broad category used to describe those who operate their own staking infrastructure. There are many configurations of solo staker, from a person at home running their own node, to operators providing smaller-scale node operator services for friends and family, to groups operating a node together by "squad staking" with distributed validator technology (DVT). Solo stakers are critical to Etheruem's health: as the network's least correlated operators, they help to distribute the control of the network among a larger number of independent participants, reducing the risk of centralization and supporting resilience with widely-distributed operations.

As participation has expanded, stakers continue to prefer staking through specialized operators. Since staking's earliest days, many stakers have outsourced the operation of validator node infrastructure to staking-as-a-service (StaaS) providers; Messari estimated that the top 15 StaaS providers had over $43B in assets staked across 12 proof of stake (PoS) networks in December 2021, well before the merge to PoS made staking possible on Ethereum. Today, just one of those operators—Figment, who had an estimated $1.8B in tokens staked at the time of Messari's 2021 report—reported over $15B in assets staked on their platform at the end of Q1 2024.

The continued trend of delegating one's ETH staking operations is not negative; the broad development of teams with dedicated expertise in node operations has improved the technical efficiency of staking participation and expanded Ethereum's global accessibility and adoption. Staking technologies like liquid staking tokens (LSTs) have crucially enabled stakers to participate in Ethereum using fewer than 32 ETH (representing a globally-inaccessible $98K USD at the time of writing), while today, DVT is beginning to enable staking fewer than 8 ETH in an evenly-distributed 4-operator cluster.

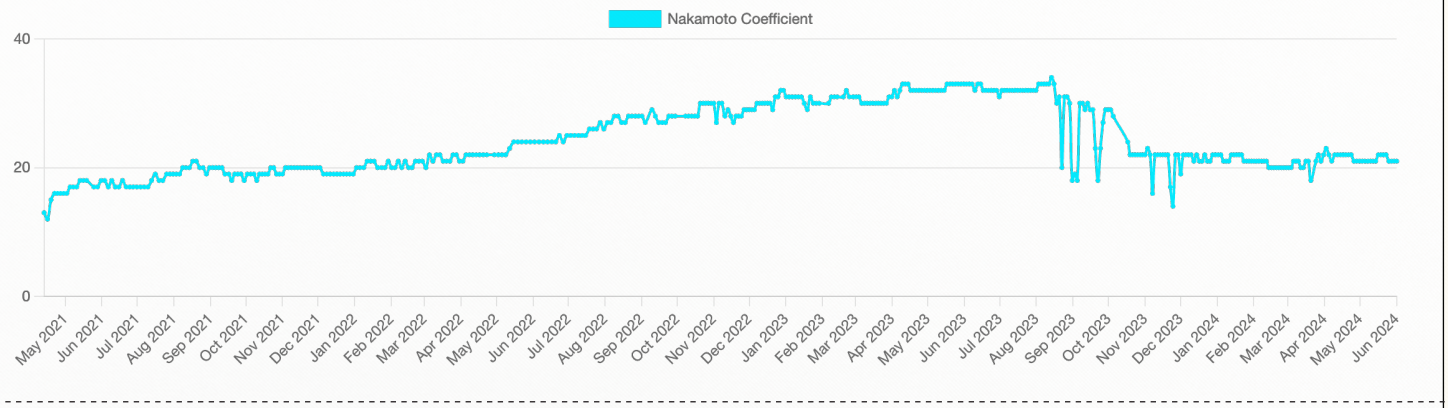
However, the prominence of staking through these kinds of solutions means that stakers and operators need to be conscious of preserving diversity and resilience while participating in non-solo staking solutions to protect themselves (and the network).

IN TANDEM, ETHEREUM'S CORRELATION RISKS BECOME MORE PROMINENT

Today's crypto-native risk-seeking behavior centers around EigenLayer's restaking innovation, which opens the potential for applications to use shared security from ETH. This innovation further increases the importance of the underlying validators' resilience against any single point of failure, as a slashing event at the core ETH staking level could lead to waterfall effects on restakers, too.

Similarly, when it comes to institutional trends, the possibility of staked ETF assets has the potential to hugely increase demand and ownership of ETH. While staking was not included in the recent US approvals of spot ETH ETFs, many predict that staking will be included in the future. Due to their potential size, staking-backed ETFs could also present the risk of centralizing staking participation—if onboarded via staking solutions that do not sufficiently diversify, including across operators. Solana provides one historical example of how quickly network diversity can shift given the institutional-scale onboarding of stake to a single-operator provider.

The Nakamoto Coefficient represents the lowest number of node operators in a PoS network who would need to collude to censor the network's transactions. In Fall of 2023, the post-collapse estate management for FTX moved the estate's SOL to be staked with a single provider, likely with the aims of simplicity and compliance. However, doing so reduced the Solana networks' Nakamoto Coefficient from 30 to 20, dramatically impacting the global network's resilience and decentralization in one swoop.



After climbing past 30, Solana's Nakamoto Coefficient fell dramatically following the reallocation of FTX stake to a single node operator.

SOURCE: SOLANA COMPASS

Any centralization of node operations like this, be it one estate transitioning to a single node operator or one ETP onboarding to a single node operator, increases the risk of a network-wide slashing incident. With over 27% of ETH supply currently staked, a correlated slashing incident could have catastrophic effects on all ETH holders, affecting both the economic and social security of Ethereum. In today's landscape correlation risk overall is also paramount beyond correlated slashing. Vectors such as client diversity represent other risks to Ethereum's long-term operational health, like a forked chain or loss of finality.

The growth of ETH staking has brought challenges for stakers, who have had to make decisions about who to stake with, and node operators, who have to decide how to scale their operations. Delegators and operators need to be conscious to protect themselves (and others) by preserving diversity and resilience while using shared infrastructure. For any staker delegating to a single, large, undiversified operator, correlated slashing represents a very rare, but very real, risk, with large implications.

2.

A Risk-Adjusted Lens

Let's examine Ethereum's current risk landscape and current trends, to highlight today's options for risk mitigation and empower educated participation in staking through a risk-adjusted lens.

The Risk Landscape: Client Diversity, and the Supermajority Client Concern

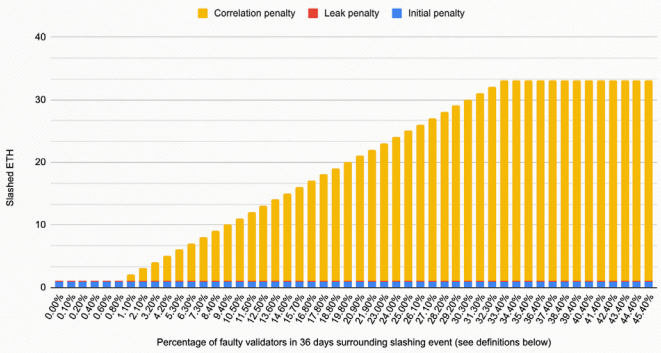
Almost ten years after Ethereum was created, validators are running on five consensus clients and four execution clients. Client diversity has improved remarkably over the past two years, but there is still a long way to go. Issues persist: a 2023 Prysm issue resulted in a lack of finalization on the chain, while a Nethermind issue earlier this year took down 8% of validators.

Thankfully, today's level of client diversity has prevented Ethereum from suffering any chain outages that are all too common on single-client chains like Solana, but risks remain. A client with a critical bug could cause validators running that client to unintentionally violate consensus rules, for example, by proposing or attesting incorrect or conflicting information about the state of the blockchain.

If this buggy client was being run by a two-thirds supermajority of validators on the network, they could finalize an invalid block, thereby causing an invalid chain fork to become Ethereum's canonical chain. For many, this potential represents the "real" doomsday scenario of correlation risk. With a majority of stake now finalizing an invalid chain, what happens next? Manual intervention by the ecosystem at large to roll back the invalid fork, or another splitting of the ecosystem such as ETH vs. ETH Classic?

If the nature of the client bug constitutes a slashable offense, a correlated slashing incident would take place, resulting in much higher penalties than the usual 1 ETH per validator slashing penalty. Correlated slashing penalties begin when 1% of validators on the network are being slashed, with the penalty constituting the validator's entire stake if over 33% of the network is being slashed. If 25% of the network's validators are slashed, due to a bug in one of the four execution clients for example, each validator loses 25 ETH. That penalty is equal to 78% of each validator's capital, and would result in 19.5% of all staked ETH being burned.

Slashing penalty per validator vs share of the network slashed in the 36 days surrounding the slashing event (with an empty exit queue)



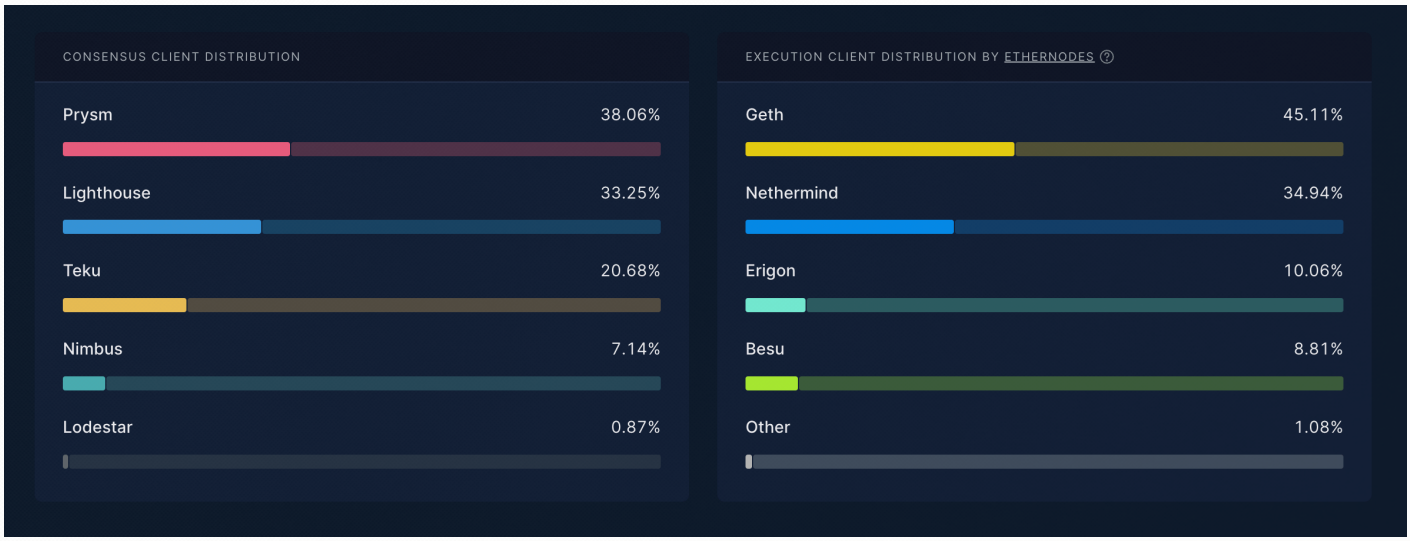
The more ETH that is slashed on Ethereum around the same time, the higher the correlation penalty.

SOURCE: KILN - USEFUL ETHEREUM SPREADSHEETS: CORRELATED SLASHING PENALTIES

This raises the question: what is the optimal endgame for client diversity? Is the goal for Ethereum's validators to be perfectly split between the four execution clients, and the five consensus clients? Even in this "perfect" world of client diversity, the risks are still huge. Is the solution to perpetually hop from one client to another, always running the least-used client? For node operators, there is a need to balance client diversity with acceptable validator performance standards, as newer, less adopted clients are by definition the least well-tested, and probably the most likely to experience issues.

It's clear that Ethereum needs even more client implementations, but this can take years. Nethermind, for example, is finally having "its moment", with almost one-fourth of staking running its client, but this level of performance and popularity was years in the making: Nethermind was founded in 2017 after a bug in Parity's wallet product (and a resulting dispute) halted operations of the Parity client, leaving Geth as the only execution client remaining. By early 2022, four years into their journey, Nethermind was operating on just over 1% of the network's nodes, and by January of 2024 it was operating on only 8.18%.

Client Distribution



Distribution of validators across consensus and execution clients on Ethereum as of April 30 2024.

SOURCE: RATED NETWORK EXPLORER

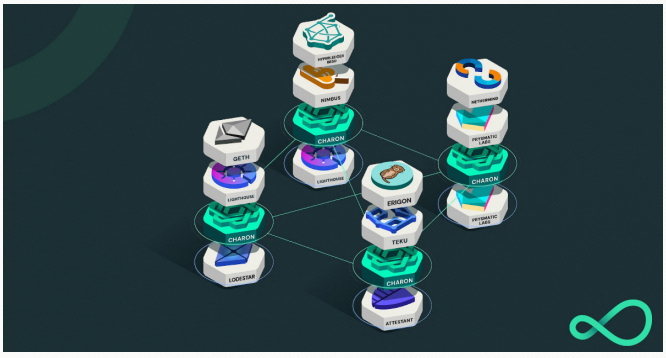
MITIGATING CLIENT RISKS

There are multiple options for mitigating the client risk when staking. Today, the core, baseline approach to improving resilience against client failures is to stake through solutions that require multi-client validator operations. By ensuring that validators are run on a variety of consensus and execution clients, a liquid staking protocol (LSP) or StaaS provider can maintain resilience against client bugs and reduce the risk of large-scale slashing incidents.

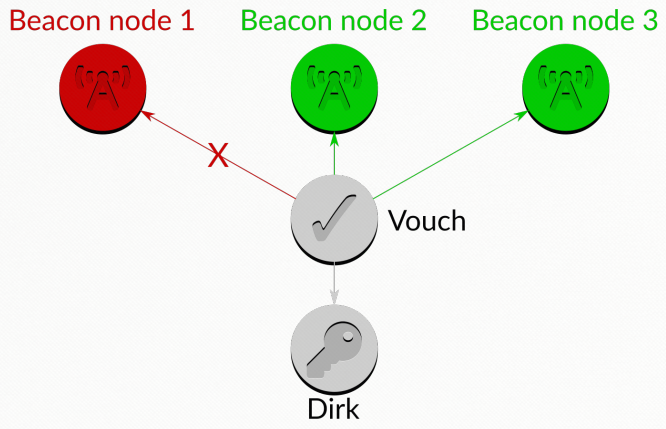
Even better than diversifying a set of validators across several clients is to use technologies like DVT or Vouch to enable multi-client validator setups. With a distributed validator (DV), multiple nodes form a "cluster" and run the validator(s) together. The cluster's nodes can be geographically distributed and run unique client stacks. If a client issue causes a node to go offline, the validator continues to operate as long as a 2/3 threshold of nodes are online.

Similarly, Attestant's Vouch is a validator client that can choose from data provided by multiple consensus layer clients (beacon nodes) in order to propose a block. An issue with a single client would therefore not cause an issue with the validator, which can continue to operate with only two normally-functioning consensus clients.

A native multi-client solution like this is the only way to totally protect a validator against issues with a single client. And with the added benefit of fault tolerance and reduced key management risks, it is clear to see why Vouch and DVT are becoming ever-more popular with professional node operators. As a developing technology DVT is being actively tested by operators from solo stakers to enterprise operators, with potential risks and limitations being identified; Obol's Charon Threat Model provides an overview of distributed validator middleware security considerations.



Obol's Charon DV Middleware enables a validator to be run by multiple nodes operating as a cluster. Each node can use its own unique client stack, reducing client risk.



Attestant's Vouch validator client can choose from data provided by three beacon nodes before it performs validator duties, reducing client risk from any one node.

Technologies like DVT and Attestant's Dirk can also reduce slashing risk. In both solutions, each node runs a key share instead of an entire validator key, requiring a two-thirds threshold of key shares to aggregate validator duties like voting. If a single key share or minority of key shares is accidentally duplicated across two nodes, there is no risk of slashing, as a two-thirds majority is required.

Large-scale node operators have a responsibility to mitigate the risk of client diversity. They also have a responsibility to balance their client choices and middleware configurations with adequate staking performance and stability. Given the reality of today's client diversity risk landscape, key mitigation considerations include whether an operator is making use of multi-client infrastructure.

TL;DR

Client Correlation Risks, and Mitigations

RISKS

- Calls for the staking industry to diversify have led to the more evenly-distributed adoption and use of five consensus clients and four execution clients. But with large percentages still using individual clients, risks remain.
- A client with a critical bug could cause validators running that client to unintentionally violate consensus rules, while less critical client bugs have been known to cause more commonplace slashing incidents.
- A buggy client being run by a supermajority of the network's validator nodes could lead to wide-scale correlated incidents, from a correlated slashing event to a fork in the canonical chain.
- Developing and refining performant, widely-adopted client technologies is an intensive, long-term undertaking, which limits the landscape of viable client options.

MITIGATIONS

- Stake through solutions that require multi-client infrastructure.
- Ensure that validators are run on a variety of consensus and execution clients, including diverse backup client configurations, with distribution from any supermajority clients .
- Use technologies like DVT or Vouch to enable multi-client validator setups, which can provide diversity even in a limited client landscape.

The Risk Landscape: Operator Diversity

The old wisdom "don't keep your eggs in one basket" is easily applied to staking. Staking assets with a single node operator is common, but it exposes the entirety of staked assets to risks associated with that node operator, including downtime and slashing.

Beyond the risk to the staker, a critical lack of operator diversity could undermine Ethereum's sybil resistance, making it easier for an adversary to coordinate attacks or cause network disruptions. A recent paper published by Eric Budish, Andrew Lewis-Pye, and Tim Roughgarden, The Economic Limits of Permissionless Consensus, provides an argument that Ethereum's economic security has improved post-merge to Pos, "formalizes the potential security benefits of proof-of-stake sybil-resistance coupled with slashing," and provides mathematical justifications for a number of Ethereum's staking design decisions. While we'll cover this paper's findings more in the "what's next" section below, it highlights the importance of sybil resistance for maintaining economic security in its discussion of achieving "EAAC (expensive to attack in the absence of collapse)." The researchers emphasize the crucial need for a diverse and robust validator set to prevent any single entity from gaining disproportionate control.

Manually staking assets across multiple operators is not always easy, often representing an accounting headache. As discussed in the "The stakes keep increasing" section above, the drop in the Solana Network's Nakamoto Coefficient following the transition of the post-collapse-FTX-estate's SOL to one node operator's management provides a warning tale of the centralization that an institution onboarding to one sole provider can bring.

Given the growing demand from investors seeking to integrate ETH into their portfolios following the launch of spot BTC ETFs in January 2024, the recent approval of spot ETH ETFs in the US was long-awaited. Globally, spot ETH ETFs are widely expected to transition at least part of their portfolios to be staking-backed; as Matt Leisinger wrote in an op-ed on liquid staking and spot ETH ETFs:

> "ONCE THE INSTITUTIONAL FLOODGATES OPEN WITH A SPOT ETH ETF APPROVAL, THERE WILL BE A SCRAMBLE TO PARTICIPATE. SIMPLY HOLDING ETH WON'T BE ENOUGH; THESE INSTITUTIONS WILL PIVOT TO STAKING, LOOKING TO GIVE THEIR INVESTORS A HIGHER RETURN ON THEIR ETF HOLDINGS (ESPECIALLY ON A DEFLATIONARY ASSET)."
>
> —MATT LEISINGER, CEO & CO-FOUNDER AT ALLUVIAL

And as Financial Times' Philip Stafford wrote in coverage of the US spot ETH ETF approval:

> "... ETF PROVIDERS ARE MISSING OUT.AS CCDATA POINTED OUT, PURCHASING 1,000 ETHER ON JANUARY 1 LAST YEAR WOULD HAVE TURNED $1.2MN INTO $3.66MN. STAKING THE SAME AMOUNT WOULD HAVE NETTED YOU $3.87MN, A GAIN OF $217,000 OR JUST UNDER 6 PER CENT MORE. IF ETF ISSUERS CANNOT GET THE YIELD, THEN THEY MAY HAVE TO CHARGE HIGHER FEES TO CUSTOMERS TO COMPENSATE. ... AGAINST THIS BACKGROUND IT'S NOT HARD TO IMAGINE THAT WALL STREET WILL BE LOBBYING TO ERASE THE BAN. MORE THAN THAT, IT MAY NO LONGER BE IN THE SEC'S BEST INTEREST TO BE DEFENSIVE ON ETHER."
>
> —PHILIP STAFFORD, FINANCIAL TIMES

While we shouldn't suppose that the participation dynamics of a spot ETH ETF in the US markets will mirror that of the spot BTC ETFs, the numbers can be used to provide an illustrative example.

- Blackrock's $IBIT spot BTC ETF saw $13.9 billion in net inflows from Jan 11 - Mar 28, 2024.
- As of May 2024, at an ETH price of just over $3k and current staked supply of 27.16%, if a spot ETH ETF were to see similar inflows in dollar terms, and stake 50% of that ETH, the ETF's staked ETH would represent ~6.2% of all staked ETH.
- If this were to be staked via a single node operator, and that operator were to replicate an issue causing a slashable offense across all of the ETF's nodes, considering the correlation penalty of 5 ETH a total of ~6.07 ETH would be slashed from each of the ETF's validators, resulting in a total loss for the fund of ~438,412 ETH, the equivalent of ~$1.38B burned.

While this example is only a long-tail risk, it illustrates the need for institutions to carefully evaluate operator diversity when onboarding to ETH staking solutions. This example would represent one of Ethereum's worst days, and would likely have waterfall effects including an impact to ETH's market value, therefore impacting every ETH holder and participant, even those who aren't staking, and who weren't directly affected by the event.

In the event that an error in node configuration is replicated across one company's infrastructure, the fallout can happen quickly. In November 2023, Bitcoin Suisse was slashed as a new segment of their nodes were activating on the network, a result of a mistake replicated across their infrastructure configuration. While their clients' finances were made whole, and the incident was localized to their own single-operator infrastructure, the equivalent of over $200,000 in ETH was burned in less than two hours as nearly 100 of their nodes were slashed and involuntarily sent to the exit queue.

For optimal diversity, individual stakers should evaluate their ability to operate home nodes, and may explore options like DVT to "squad stake" for improved infrastructural resilience while enabling staking in amounts fewer than 32 ETH. Resources like the ETHStaker Community provide information to educate solo stakers on critical diversity best-practices, including evergreen resources on infrastructure management and initiatives such as the Obol x ETHStaker DVT Home Staker Program.

While larger-scale participants may feel that operator diversity brings an additional burden of managing staking accounts across multiple providers, many staking solutions exist that diversify across multiple node operators to reduce correlation risks. One solution is to stake through a liquid staking protocol which distributes staked tokens through multiple providers within its active set. This allows stakers to interface with one solution, while keeping their "eggs out of one basket." Composable staking solutions, including LSTs like Liquid Collective's LsETH, can help provide stakers with multiple entry-and-exit points, reducing the risk of any one node operator presenting a single point of failure.

Regardless of the provider(s) used, care should be taken to ensure that they have anti-slashing mitigations in place which prevent double-signing, both on the client level and at the key/validator level. The use of key-sharding technologies like DVT or key managers like Attestant's Dirk or Web3Signer can protect against slashing, in addition to clients which implement their own anti-slashing checks.

TL;DR

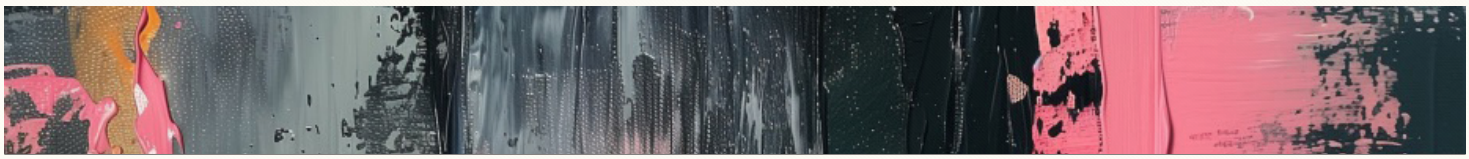
Operator Correlation Risks, and Mitigations

RISKS

- Staking assets with a single node operator is common, but it exposes the entirety of staked assets to risks associated with that node operator, including downtime and slashing.
- A lack of operator diversity can undermine Sybil Resistance if a large portion of Ethereum's validator set is controlled or influenced by a few entities, making it easier for an adversary to coordinate attacks or cause network disruptions.
- While the percentage of circulating ETH currently staked is at all-time highs, the scale of institutional onboarding can still represent a significant percentage of a network's validators.

MITIGATIONS

- Large-scale stakers should distribute their staking operations across multiple node operators. While staking across a diverse set of operators is an important consideration for any staker, the scale of institutional stake onboarding amplifies the importance of diversifying across operators.
- Individual stakers should evaluate their ability to operate home nodes, and may explore options like DVT to "squad stake" for improved infrastructural resilience. while enabling staking in amounts fewer than 32 ETH.

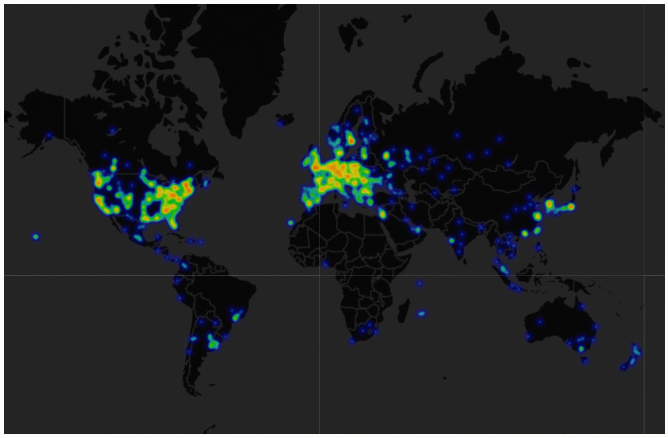


The Risk Landscape: Cloud and Regional Diversity

In addition to reducing reliance on any single staking provider, care should be taken to reduce reliance on any single cloud or hardware provider, or any particular geographic region or country.

Geographic diversity and cloud diversity both carry correlated outage risks. For regional diversity, the risks may be more aligned with other considerations beyond slashing, such as latency in timing games affecting SRR, but the regional centralization also carries correlation risk of a regional "blackout" of a cloud provider or critical internet infrastructure.

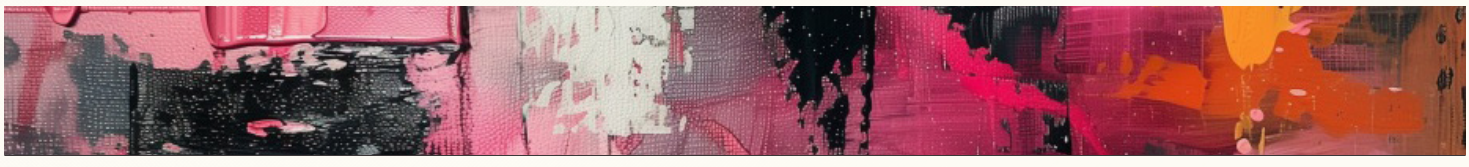
Geographic diversity is also necessary to avoid problems stemming from natural disasters or wars. Japan's 2011 earthquake and tsunami tested the resilience of data centers and the entire power grid. At the outbreak of the Ukrainian war, the largest Solana validator was the Ukrainian company Everstake. Thankfully, they were prepared, but these scenarios are not always predictable. Countries can also change their policies toward crypto, with the most extreme example probably being the 2021 ban on Bitcoin mining in China, which caused Bitcoin's hashrate to fall by a 50% in two weeks as miners were forced to relocate.



Ethereum's validators' geographic distribution as of May 10, 2024.

SOURCE: ETH2 NODEWATCH

There's also no guarantee that hardware or cloud providers remain supportive of crypto: in 2022, the popular hardware provider Hetzner took offline over 20% of Solana's validators, blocking access to those customers' servers, forcing them to quickly migrate their validators elsewhere.

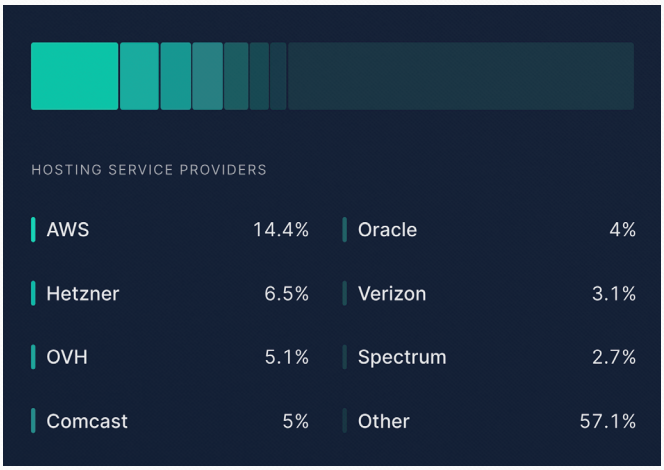


Providers are also not immune to downtime, with AWS suffering three outages in December 2021 alone. Too many validators running on a single hosting provider can put the entire network's ability to achieve consensus into question.

A lack of cloud provider/region diversity adds a layer of single-operator risk as well. You don't have to be familiar with staking's correlation risks to be keenly aware of how regional cloud outages can affect business continuity.

Perennial outages in Amazon Web Services' (AWS) regional data centers make headlines for taking businesses offline that aren't diversified across regions and providers, from financial institution T. Rowe Price's users losing access to their personal investing and trading accounts, to New York's Metropolitan Transportation Authority losing its service disruption reporting, to the Taco Bell and Burger King ordering apps going down.

Host Distribution



Ethereum validators' cloud distribution as of May 6, 2024.

SOURCE: RATED

As new validators are spun up by a liquid staking protocol or other staking provider, validator keys can be provisioned across clouds and regions to mitigate correlation risks. Large-scale participants should also consider factors like failover protection to ensure that validator operators can transition smoothly to a new configuration, should they go fully offline.

DVT also has the potential to greatly improve these diversity considerations. For solo stakers (and stakers using any staking services), implementing a multi-operator staking configuration with DVT can ensure that one operator can pick up signing duties if another is to go offline due to a regional or cloud outage. Even for the largest of operators, early testing has shown that DVT configurations may be more resilient than other failover technologies.

TL;DR

Cloud and Regional Correlation Risks, and Mitigations

RISKS

- Geographic diversity and cloud diversity both carry correlated outage risks.
- Cloud outages are commonplace across web-based services, and cloud service companies can change service policies.
- Geographic diversity is also necessary to avoid problems stemming from natural disasters or wars.

MITIGATIONS

- Validator keys can be provisioned across clouds and regions to mitigate correlation risks, and failover protection configurations can ensure node operators are able to transition to a new configuration if they go fully offline.
- DVT has the potential to greatly improve cloud and regional diversity, both via multi-operator staking configurations and as a single-operator failover technology.

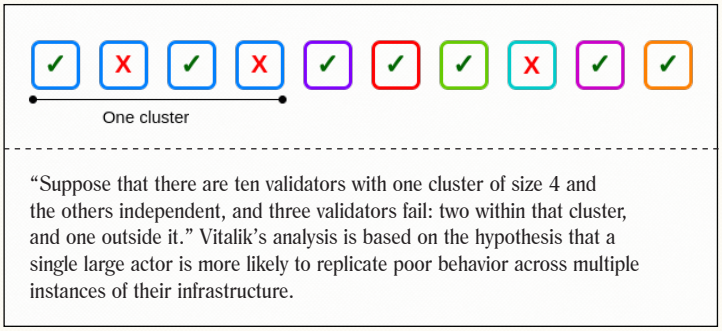
3.

What's Next

The robustness and diversity of Ethereum's infrastructure has improved significantly over previous years, but still has a ways to go. A variety of current EIPs and community discussions could shift the landscape surrounding correlation on Ethereum. If one thing appears certain in Ethereum's future, it's that the importance of correlation within staking dynamics won't fade.

Anti-Correlation Incentives

A recent proposal from Vitalik to add additional correlation penalties could provide even more momentum. Vitalik's proposal is based on the hypothesis that a single large actor is more likely to replicate poor behavior across multiple instances of their infrastructure. Vitalik's analysis of missed attestations finds that this assumption checks out, with historical data showing that "two validators in the same cluster are significantly more likely to miss attestations at the same time than two validators in different clusters."



"Suppose that there are ten validators with one cluster of size 4 and the others independent, and three validators fail: two within that cluster, and one outside it." Vitalik's analysis is based on the hypothesis that a single large actor is more likely to replicate poor behavior across multiple instances of their infrastructure.

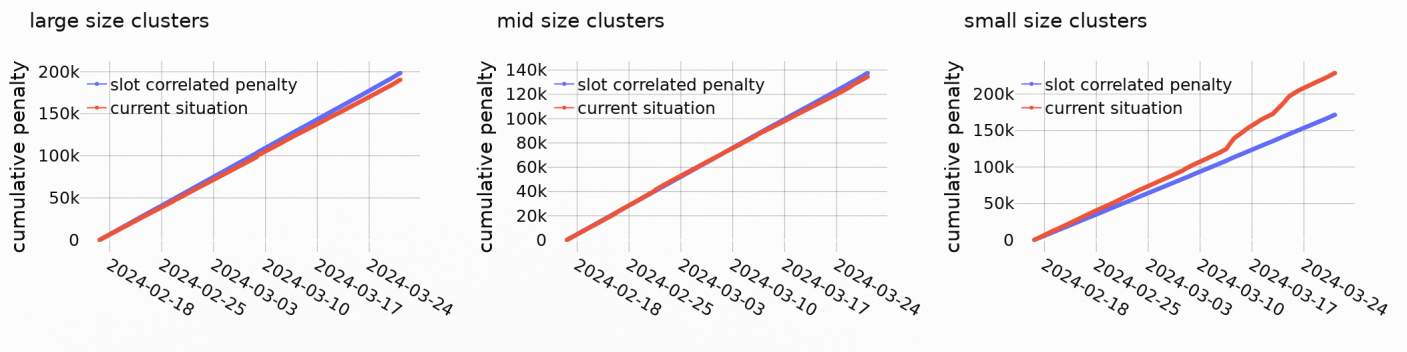
SOURCE: ETH RESEARCH

Attesting to the validity of blocks—and potentially missing that window to attest—represents a common, everyday occurrence. Every validator on Ethereum is tasked with attesting once per epoch (~6.4 minutes), which stands in contrast with the more rare, somewhat-existential risk of a widespread correlated slashing event.

Signaling a much more common mistake made at scale across many validators in a cluster, Vitalik suggests that an additional penalty for correlation within missed attestations could help adjust Ethereum's economic incentives "to favor architectural decentralization." Essentially, by providing a more constant, minor penalty that dings the staking rewards received by large operators with overly-correlated infrastructure, large-scale operators could be more actively incentivized to adequately distribute their infrastructure.

Ethereum researcher Toni Wahrstätter published a quantitative analysis of Vitalik's proposal, "Analysis on 'Correlated Attestation Penalties'." Examining onchain attestation data from February - March 2024, Wahrstätter analyzes the sum of penalties received under current network dynamics to what those penalties would have been under Vitalik's proposal.

Comparing four cohorts of operators—large, medium, small, and unidentified—Wahrstätter's analysis finds that Vitalik's proposal is effective in adding a minor increase to penalties for large-to-mid sized clusters, while the smaller cohorts, primarily representing at-home and solo-stakers, don't have to worry as much about missing the odd attestation.



The cumulative impact anti-correlation penalties would have, based on historical data.

SOURCE: ANALYSIS ON "CORRELATED ATTESTATION PENALTIES"

These results can be viewed as aligned with the dual goal to 1) incentivize large-scale operators, who maintain a larger portion of the network's nodes, to diversify their infrastructure and operate to the highest performance standards; while 2) providing a "leg up" to solo stakers to ensure distributed, smaller-scale operators aren't left behind.

Built off of Wahrstätter's analysis and Vitalik's initial proposal, Vitalik's concrete proposal for correlated attester penalties is still being actively debated and reviewed by the community as of June 5, 2024. Its implementation timeline, and whether it will be implemented at all, remains unclear. But, its positive reception signals that the implementation of further anti-correlation incentives may continue to make Ethereum's correlation risks a key consideration. With contributing factors to missed attestations including vote accuracy, latency, and other indicators of node operator performance, large-scale stakers will continue to need to evaluate a node operator's performance standards alongside risk mitigations to ensure sufficient performance and distribution.

EIP-7251: Increase the MAX_EFFECTIVE_BALANCE

EIP-7251 is currently on track to be included in Ethereum's upcoming Pectra upgrade, expected to go live in Q4 '24 – Q1 '25. While it primarily combats the potential for unsustainable network bloat due to too many validators being active on Ethereum, it will likely impact staking's correlation dynamics overall.

Today, Ethereum's MAX_EFFECTIVE_BALANCE limits the active stake of a single validator to 32 ETH. EIP-7251 proposes increasing the maximum active stake of a single validator to 2048 ETH, while retaining 32 ETH as the minimum stake required to activate a validator on the network.

With each validator node allowed to have a higher effective balance, professional node operators are expected to consolidate the number of validator nodes they are operating to reduce costs and streamline operations, leading to a decrease in the overall number of validators on Ethereum. This reduces the number of redundant validators (those operated by the same entity) and thus, the number of P2P messages and BLS signatures that need to be aggregated each epoch.

EIP-7251 also suggests removing the initial slashing penalty of 1/32 of a validator's effective balance (today, representing 1 ETH). This would aim to make consolidation even more attractive for professional node operators, as a larger effective balance would mean that an initial penalty of 1/32 would represent a larger hit (63.9 ETH, if a validator at the full 2048 ETH were to receive a slashing penalty). This section of the proposal remains under debate.

This transition to a larger effective balance will also require that relevant constants on the protocol transition from addressing a number of validators to addressing the "weight" of validators, or total amount of ETH stake impacted. For example, EIP-7251 maintains churn invariants, which were adjusted in the Dencun upgrade to similarly slow growth in Ethereum's active set, by adjusting activation and exit processes to be balance-weighted rather than count-based. Vitalik's previously-discussed proposal to add additional correlation penalties based on missed attestations is designed to accommodate EIP-7251. In the proposal's FAQ, Vitalik raises the question of whether stakers could avoid the multi-validator attestation penalty given EIP-7251's raising of MAXEB; this could be possible as operators could have fewer active validators, thus providing less of a signal for failures within the multiple validators running on one cluster. To this, Vitalik explains that within his proposal, "The proportional penalty formula would count the total amount of ETH, not number of validator IDs, so 4,000 staked ETH that acts the same way would be treated the same if it's split between 1 validator or 2 or 125."

While this consolidation of the number of validators under operation should lead to improved efficiency and reduced operational costs, it will also emphasize the importance of operating sufficiently resilient staking infrastructure at scale.

The scale of EIP-7251's impact on operational costs for large-scale node operators will vary based on their specific infrastructure configurations, but includes considerations from the administrative burden of registering and funding validator keys, to cost of running servers.

"From a professional DevOps perspective, running less infrastructure is better," explains David Turnbull, Senior SRE at Alluvial. "EIP-7251 will have a big impact on the number of validator keys that need to be funded. That brings efficiencies in administrative labor, upgrades, monitoring, and more. But one of the biggest costs for large-scale operators is cloud compute. By all means, EIP-7251 will hugely reduce this cost. For more classic infrastructure configurations, which include operating cloud-based servers, and benefiting from the elasticity of the cloud, not having to activate new validators will mean their savings will be huge. For more elegant configurations, their savings may be less, but they will still be there."

Stakers will also see the impact of these efficiencies. The increase to maximum effective balance will benefit solo stakers, who will be able to compound rewards to increase their effective balance in smaller increments than 32 ETH. EIP-7251 also introduces the ability to set a custom ceiling on a validator node's maximum effective balance, beyond which any ETH network rewards received are automatically claimed in an automatic, gasless, partial-withdrawal transaction, creating further flexibility for staking configurations. This could potentially increase network diversity by encouraging more participation from solo-stakers, who may lack an ability to "compound" rewards today due to the 32 min and max effective balance—and who may lack the technical resources to manage an ever-growing number of validators if and when they are able to compound.

Most liquid staking protocols, including Liquid Collective, already automatically stake ETH network rewards received as they reach a fungible bulk of 32 ETH—one of the key efficiencies of liquid staking technology. But EIP-7251 will still bring efficiencies to liquid staking. The administrative burden of the protocol requesting that node operators pre-generate validator keys to be ready for activation, and the administrative burden of funding new validator keys, will both be reduced. The automatic staking of rewards will be made much more efficient by removing the overhead of spinning up new infrastructure.

However, these efficiencies come with tradeoffs, which may amplify the importance of carefully evaluating correlation risks across staking post-EIP-7251.

Network diversity could be impacted by EIP-7251 if many nodes across many regions are consolidated into fewer nodes in more central regions, or many nodes running diverse clients are consolidated into fewer nodes with less client diversity, etc.

To argue that EIP-7251 is unlikely to have a negative impact on network diversity, one could make the argument that—just as with the shift for existing network dynamics, like the staking churn limit, to evaluate the weight of ETH instead of number of validators—as long as professional node operators retain their diversity targets when consolidating validator nodes, there is no measurable difference between operating 300 nodes across 3 regions vs. 3 mega-nodes across 3 regions, for example.

But the efficiency of consolidation will still have the effect of reducing the flexibility that activating new validators represents. Today, each time a node operator registers a new validator represents an opportunity to diversify their active set via that validator's configurations, potentially balancing it in a different region, etc. But, given the scale of each node reaching up to 2400 ETH post-EIP-7251, there will simply be fewer opportunities to adjust those diversity targets.

Implementing DVT may offer the solution to retain this flexibility post-EIP-7251. Because DVT enables a validator to be run by multiple nodes operating as a cluster, operators could use the diversity within each larger-sized node to access that flexibility of operating many nodes, while still benefiting from the efficiencies of operating fewer validators.

"It will be huge for solo stakers to not have to run multiple instances of infrastructure, but as a professional, I like infrastructure that's ephemeral. I want to be able to build it up and tear it down," explained Turnbull. "Things happen, security incidents happen, diversity targets and change, maintenance is required, and on a professional scale, there remains that need for flexibility when activating nodes. That's why EIP-7251 in addition to DVT is interesting to me. Adding that redundancy—the ability to tear down and bring up nodes, but within one validator—together will make operations more efficient while retaining flexibility."

While EIP-7251 is still actively under development, the design choices made to effectively accommodate the change by both node operators large and small will require careful evaluation of their impact on correlation risks, and how diversity targets and solutions like DVT can act as solutions.

The Economic Limits of Permissionless Consensus

Noted in the "operator diversity" risk landscape section above, Eric Budish, Andrew Lewis-Pye, Tim Roughgarden published a CompSci research paper in May 2024, "The Economic Limits of Permissionless Consensus." Examining Ethereum's economic security, the paper analyzes the challenges of achieving EAAC (expensive to attack in the absence of collapse) conditions without collateral damage in various network settings, and formally justifies the benefits of slashing, inactivity leaks, cooldown periods, and other elements of Ethereum's staking design.

Emphasizing the need for a diverse validator set, the paper establishes that in certain settings no protocol can guarantee that all attacks are expensive without collateral damage. This underlines the complexity and limitations of designing penalty mechanisms that are both effective and fair, while highlighting how one of PoS' core security mechanisms is the imposition of economic penalties on validators who fail to perform their duties correctly.

As technical writer Emmanuel Awosika wrote in his post about the paper, "You probably saw me duke it out with [Toghrul Maharramov, Researcher at Scroll] and [Anatoly Yakovenko, Co-Founder of Solana Labs] … on the question of "Does slashing matter?" a few weeks back. One oft-repeated argument was the lack of a formal proof that slashing meaningfully improves Ethereum's (economic) security. This paper proves that slashing *does* improve security for blockchains… Designing to slash validators is a specific/deliberate decision--it comes with tradeoffs, but also produces the outcomes the Ethereum protocol desires."

Through his paper's work of "formalizing… the common belief that the merge has increased Ethereum's economic security," it appears clear that economic disincentives, such as slashing, inactivity leak penalties, or missed attestation penalties, continue to be seen as valuable resources for maintaining Ethereum's security and performance. This paper and its economic justification may likely serve as a basis to expand anti-correlation economic disincentives within the protocol moving forward.

LEVELING THE PLAYING FIELD FOR LONG-TERM RESILIENCE

It's important to note that the thrust of these proposals and research pieces are not intended to inordinately punish those who stake through staking providers, nor to make staking economically infeasible for operators at scale. As Vitalik notes in his proposal, additional correlation penalties could help to make "... solo staking relatively more competitive by making the economic part of the incentives more balanced to favor architectural decentralization."

Similarly, Wahrstätter's analysis focuses on "leveling the playing field" to ensure that solo stakers are not left behind as technologies like liquid staking grow in popularity. As he writes in his introduction, "To strengthen decentralization, it is essential to architect mechanisms that counteract the advantages of economies of scale. Fortunately, economies of scale are intrinsically linked with correlation effects: When a single operator runs many validators on one machine and experiences downtime, all those validators are affected at once. Thus, leveraging economies of scale comes with correlation effects."

Today, there are many advantages to staking through a liquid staking protocol or other staking provider. This is why these technological solutions have been developed: to make staking more efficient and accessible. Liquid staking provides access to liquidity and capital efficiency for stakers, while staking technologies and providers overall can introduce efficiencies, like auto-staking of rewards received, smoothing of block proposal rewards, and deposit and redemption buffers.

If Ethereum is to be used for financial activities and high-value applications at scale, it must have a universal network of globally-accessible peers with constant uptime, correctness, and finality. To that end, preserving the decentralization and distribution of Ethereum is key. While economies of scale have exacerbated the potential for a correlated slashing event by a major operator, this mechanism and the attestation correlation penalty proposed by Vitalik serve to advantage the solo staker who adds incremental diversity to the network.

Awareness is Critical

With the Ethereum protocol still undergoing major upgrades and changes, stakers should stay aware of the types of risks their choices expose themselves to, and whether their stake is incrementally increasing network diversity. On the social layer, node operators can be encouraged to migrate to minority clients, or take advantage of multi-node staking technologies like Vouch or DVT to mitigate risks, and improve diversity.

In terms of staking choices, it's critical to evaluate one's participation through a risk-adjusted lens. By staking through protocols and solutions that diversify across operators, regions, clients, and clouds, actively mitigate the risk of slashing, and abide by performance and risk standards that meet one's own risk profile, stakers can help to improve and preserve Etheruem's resilience for all participants.