[

image

1745×1010 268 KB

](https://ethresear.ch/uploads/default/original/2X/7/7f4e33225d79adb3a9bf570a9a17e480c20c961f.jpeg)

Introduction

Relays in MEV-Boost are doubly-trusted stakeholders within the existing PBS landscape and yet they are not properly incentivized. People have discussed and are concerned over many ways in which validators and builders game the current system, such as slot proposal delay for timing games and colocation for latency optimization, all to maximize their MEV revenue. However, relays have been looked over as a public good who reap no benefit in being honest and trusted. Relays could easily [unbundle transactions and steal MEV]() if their rewards surpass any reputational cost and loss of future revenue. But so far, most relays don't dare to collude – they may prefer to just quit relaying instead. Hence why I have [shed some]() light on relays' importance and explored the incentive dynamic across PBS with the hope of devising a sustainable plan for relay incentivization. With this piece, I would like to point out a new direction for thinking about this relay incentive problem from the lens of missed slots and propose approaches for profitable competitive and cooperative relaying.

Problems with existing proposals

Largely in two parts:

1. Incentive misalignment across validators and/or builders:

2. Relay centralization – builders are incentivized to send bids to one relay only.

Most of the proposals to date (check the[here]() for more details) are trying to carve up a piece of builders' bids for relays, which means that builders and/or validators are losing potential future profit by accommodating relays. Implementing such proposals will mean that you have to ask at least one of the two parties to make some compromises and requires a large amount of off-chain coordination work, which is hard to execute if not incentive-aligned.

Alternatively, building a competitive relay market could enable a path to monetization by hyper-optimizing relay latency. However, the current proposal of competitive relays captures values by out-competing other relays in latency and aggregating more bids. For context, here is an explainer diagram for competitive relay.

[

1280×1157 98 KB

](https://ethresear.ch/uploads/default/original/2X/d/dde11e6ef4b47f4a7ed22d8e6c05c6e9f1e30f72.png)

Figure 1

. Diagram illustrating the bid flow among competitive relays with rebates

In this diagram, winning builder 4 sends the bid to relay 3 who then rebates a maximum of 0.029 ETH

back to the builder 4. The bid transfer of the block is modified such that the relay's bid amount and the rebate can be adjusted accordingly. Builder 4 will receive back the rebates, the delta minus the relay fee. And if there are competing relays, a market-drive relay fee reduction is expected. As such, competitive relay may lead to relay centralization where builders are incentivized to send their bid exclusively to a single relay that gives the most rebates with the best latency. If we were to extrapolate the competitive relay market, we will likely see a future where each block will be broadcasted by a single winning relay, instead of the more collaborative one that we see today. This raises a few questions:

1. How bad would the slot miss rate be if we assume all the blocks to be relayed by a single winning relay under the competitive relay market?

2. What is the cost of missed slots and what would that be for competitive relays?

3. What alternatives there are to improve the miss rate while better incentivizing relays?

Analysis of Slot Miss Rate via Single Relay Broadcasting

Existing competitive relay proposals attempt to double down on the latency game where they charge the delta between the highest bid and the second-highest bid from other relays

. The builder will receive up to the delta amount of rebates, and the relay subsequently charges a % fee. Most builders today send their bids and the block to multiple relays with the hope of maximizing the block inclusion. However, with bid rebates offered via competitive winning relay, builders now have an incentive not to submit their bids to other relays. Therefore, bids will converge over one or two relays that most builders trust. One of the immediate problems that may arise from such centralization is the higher chance of missed slots

for the network. Since there are fewer relays to broadcast the block payload when called getPayload from proposers, there will be a higher slot miss rate. However, whether it's actually higher, how much higher, and how it correlates to the number of relays that propagate the payload is yet to be investigated thoroughly. To provide more context in an attempt to answer those questions, here are the results based on analyzing 30 days of relay data (across Nov ~ Dec 2023).

To determine whether single-relay broadcasting, where only one relay broadcasts the payload, results in more missed slots, I have drawn Figure 2. which shows the % of slot miss rate within its corresponding relay type, single vs cooperative. The results here show that single relay broadcasting results in 0.215%

of missed slots which is 3.18 times

higher than that of cooperative relay broadcasting, which is 0.0675%

. The chart also indicates that single relay broadcasting constitutes about 74.6%

of the total missed slot across the network, while the other 25.4%

comes from cooperative payload broadcasting. Essentially 3 out of 4 missed slots are due to single relay broadcasting. We can observe that the slot miss rate is significantly higher for single-relay broadcasting.

Figure 2.

Corresponding % of slot miss rate with payload broadcasted via single VS multiple relays (cooperative)

So far, I have grouped all cooperatives relaying under one bracket. But in reality, there could be 2, 3, or more relays participating in broadcasting, all of which may result in different slot miss rates. Hence, I have also categorized the slot miss rate under different numbers of broadcasting relays in Figure 3. This diagram shows that the slot miss rate decreases exponentially with an increasing number of broadcasting relays, and there are 0 missed slots with 5 or more relays broadcasting the payload

.

Figure 3

. Corresponding % of slot miss rate with different numbers of broadcasting relays

Some relays may have different network connectivity with diverse sets of registered validators, hence varied performance in payload broadcasting. To understand how different relays perform, Figure 4 categorizes the slot miss rate per relay during its single-relay broadcasting. The results suggest significant discrepancies in slot miss rate across relays where the highest miss rate is 1.52%

(manifold) while the lowest is 0%

(aestus). Other major relays like Flashbots, Ultra sound, and bloXroute have a miss rate of around 0.2%

. It's worth noting that there is also a significant discrepancy across the sampling size of the payload delivered for each relay. Smaller relays like Eden and Manifold tend to have higher miss rates due to the lower number of total payloads delivered as noted in Table 1. Additionally, Aestus and Agnostic have only 4~6%

of their payloads broadcasted without any cooperative relays, which is considerably lower than others. This explains why those two have a relatively lower slot miss rate in Figure 4.

Relay

No. of slot broadcasted (total)

No. of slot broadcasted as the single relay (% of total broadcasted payloads per relay)

No. of slot missed when broadcasted with a single relay

Slot miss rate with single relay broadcasting (%)

Flashbots	61116	28126 (46%)	63	0.22
bloXroute_Max	49738	19930 (40%)	38	0.19
bloXroute_Reg	67214	16117 (24%)	54	0.34
Ultra sound	99428	28418 (29%)	45	0.16
Agnostic	63475	2316 (4%)	3	0.13
Aestus	8800	486 (6%)	0	0
Eden	321	265 (83%)	2	0.75
Manifold				

231

66 (29%)

1

1.52

Total

350323

95724

206 (Total missed slot: 359)

Table 1

. Slot counts on delivered and missed payloads for each relay

Figure 4

. % of missed slots per relay without any cooperative relay

Few observations have become clear after going through the data.

1. Single relay broadcasting has an exponentially higher slot miss rate than cooperative relay broadcasting.

2. ¾ of the missed slots today come from single relay broadcasting.

3. Relays have varied slot miss rates. Well-connected relays (Flashbots, bloXroute, Ultra sound) have an average of 0.2% and less-connected relays 4-7x higher miss rate. Also, note that some less-connected relays can still achieve a low (in this case 0) miss rate by always participating in cooperative relaying.

It requires more data to attribute the exact cause for the higher miss rate among single relay broadcasting. What we are seeing here is observations that show some correlations between single relay broadcasting and missed slots. But does not definitively provide the cause nor the attributions as to why such a correlation exists.

But here are some possible reasons that I see so far.

1. Network topology:

Some proposers might be positioned quite far from a relay. The roundtrip of sending getPayload and receiving and broadcasting that payloads take additional time that results in a higher chance of validators missing the attestations.

1. Validator registrations:

Some solo validators may only register with a few relays and have a narrower view of the incoming bids/builders. As such, it results in a higher miss rate when proposing due to other relays not receiving the getPayload calls.

1. Relay Operation:

Relay may have downtime, bugs, and bad network connectivity. This could be remediated via low-latency propagation networks like bloXroute's BDN or Chainbound's Fiber Network. But relays are not well-incentivized at the end of the day ;((

1. Builder's bid bias:

Certain builders could be sending bids only to a single relay, which contributes to a higher miss rate.

1. Timing game:

Proposer might be delaying the slot with the hope of receiving higher value blocks.

Now that we have a better picture of the slot miss rate, I can also extrapolate what the miss rate under competitive relay will look like. Assuming all the bid fragments and split into corresponding relays of their choice to maximize the relay rebates, it will have a slot miss rate that ranges from ~0.13% to ~1.52% (Figure 4 & Table 1). Given over 90% of the slots are broadcasted by major relays (Flashbots, Ultra sound, bloXroute), it's safe to say that the current projected slot miss rate under a competitive market will be around 0.2%. With this number in mind, we can quantify the cost of missed slots and explore ways to incentivize cooperative relaying to minimize the slot miss rate.

Note: The number here does not account for missed slots induced by the timing game. Likely, the slot miss rate will be higher as the proposer timing game becomes more prevalent. Also, the number I have projected here is based on the Nov-Dec data and may not reflect the real missed slot since the relay and validator situations are constantly changing.

Cost of competitive relays = Cost of proposers missing slots

Currently, there are no penalties for missed slots for proposers and validators. The main "cost" here would be the opportunity cost of not receiving the rewards issuance. Here is the current rewards issuance scheme:

[

738×340 15.7 KB

](https://ethresear.ch/uploads/default/original/2X/6/684f6c196a8736de2222fa005436bdc58f3317cd.png)

Table 2

. Rewards issuance weighting by eth2book

The rewards are calculated per epoch where each epoch consists of 32 slots. The total available rewards issuance across all validators N are:

$$NR_{A} + 32(512R_{Y} + R_{Ap} + R_{Yp}) = Tb$$

where N

is the number of validators, R_{A}

is the rewards for a single attestation, R_{Y}

is the reward for a single sync committee contribution, R_{Ap}

is the reward for a block proposal due to attestations, R_{Yp}

is the reward for a block proposal due to sync committee contributions, and Tb is the maximum issuance per epoch in Gwei. A proposer missing a slot results in the loss of $R_{Ap} + R_{Yp}$

rewards per slot.

Let's take the number of validators as 900,000 and calculate the total proposer rewards per slot. With Eq 5

from the Appendix

, it will be around 0.042 ETH

per slot in the issuance rewards with an additional few thousand Gwei of priority fees and builder's bid rewards from MEV-Boost that could range from a few ETHs to tens of ETHs (an average bid of 0.14 ETH). The total opportunity cost on average will sum up to be around 0.1442ETH per missed slot

.

The current average slot miss rate overall is:

$$\text{ Total missed slot / Total slot } = \frac{359}{350323} \times 100 = 0.10 \text{%}$$

out of which 27.3% of the payloads are delivered via a single relay. Based on the current miss rate, the cost of missed slots per day is:

$$\text{Cost of missed slot per day} = 7200 \times 0.0010 \times 0.144$$

$$= 1.0 \text{ ETH}$$

Under competitive relay with single relay broadcasting, I assume the overall slot miss rate tends towards around 0.2%. As such the cost of missed slot per day will be:

$$\text{Cost of missed slot per day} = 7200 \times 0.0020 \times 0.144$$

$$= 2.0 \text{ ETH}$$

$$\text{Cost of missed slot per year} = 2 \times 365$$

$$= 730 \text{ ETH}$$

While this number is a rough estimate, an extra few hundred ETH per year could certainly cover the cost of multiple relay

operations while incentivizing cooperative relaying. In the future where relay centralization happens among competitive relays, the cost of missing slots will double for proposers. This is where cooperative relaying comes to the rescue.

If proposers have 4 or more relays cooperatively broadcasting the blocks, there would be almost no missed slot

, reducing the missed slot cost for proposers. Note that the costs we calculated are for proposers, hence should not conflate it as a network cost.

Missed slot = Relay Incentives??

Missing slots result in an opportunity cost, which sets the upper bound for what relays could be feasibly paid to prevent such losses. Paying relays in this manner functions similarly to insurance. For solo validators especially, missing a slot incurs higher opportunity cost given the lower chance of proposing compared to large operators. To help reduce missed slots,I envision a viable path for relays to become profitable by offering Broadcasting as a Service (BaaS). To achieve this, however, we need a relay mesh network

where a winning competitive relay could rapidly distribute the payloads post-commitment so that other relays can help propagate the payloads and minimize the missed slot.

Direct Channel Relay Communication Network (Relay Mesh Network):

To ensure that a competitive winning relay can quickly distribute block payloads and corresponding fees for other relays participating in BaaS, it is necessary to establish dedicated, low-latency direct communication channels between relays. These channels need to be significantly faster than the standard devp2p. One potential approach could be to set up a modified TCP protocol or utilize UDP for improved latency between known relays (reference here for more). By creating a low-latency relay communication network, we can enhance relay incentivization in two ways: firstly, by incentivizing the competitive relay through bid rebates, and secondly, by distributing a portion of the fees across other relays for cooperative broadcasting. There are existing low-latency networks offered, such as BDN by bloXroute that service both subscribing relays, traders/searchers, and validators. While the purpose is different from the relay mesh network, the goal of delivering trades/bids/payload fast is the same.

Enshrining Missed Slot Penalties (a nice-to-have minimal viable enshrinement for relay incentivization):

I originally started writing this piece to explore possible enshrinement for enabling relay incentivization. However, I realized after looking through data on missed slots that there is a relatively simple minimal viable enshrinement that will increase the relay profitability: missed slot penalty on proposers. Max has also prompted this idea nicely in his tweet here.

[

599×511 32.4 KB

](https://ethresear.ch/uploads/default/original/2X/3/37422bd460e822fb64fc931b4662d6bed9f39042.png)

Max's post on missed slot penalties on the proposer timing game

Although his comment on missed slot penalty was aimed at fixing theproposer timing game, I believe it is beneficial for relays as well. There is a clear opportunity cost for missing a slot as a proposer, which sets the max bound for the fee that relays could earn. However, this bound could be further increased if there is a missed slot penalty on top of it. Assuming we take 15M base fee (100 Gwei) for the penalty as Max suggested, missing slots would cost additional 1,500,000,000 Gwei = 1.5 ETH which increases the cost of missed slots per day from about 1 ETH to 11.8 ETH per day. If under competitive relay, this number could increase up to 23.7 ETH per day. This is a very significant increase in cost but certainly should be taken with a grain of salt given the assumptions around penalty calculations and constant slot miss rate. Nonetheless, increasing the cost of missing slots will increase the value of BaaS and hence enable better cooperative relay incentivization on top of competitive relay rebates.

Challenges ahead With Missed Slot Penalties

One major concern around missed slot penalties is raising the barrier of entry and disproportionate penalties to solo stakers. Unlike large operators, solo stakers have potentially sub-optimal network connectivity, and less capital/manpower to maintain the validators. According to Figure 5 & 6, which look at missed slots from a validator angle, solo stakers roughly have a 4~6% miss rate while professional operators like Figments have 0.2~0.4% miss rate.

Figure 5

. % of missed slots for solo stakers fromtiming.pics by Toni

Figure 6

. % of missed slots for Figment from timing.pics by Toni

Based on some napkin math, blindly imposing every validator with missed slot penalties will punish the solo stakers magnitudes more than professional stakers like figments. This will likely incentivize solo stakers to migrate to staking pools in an attempt to minimize penalties. Hence, if the missed slot penalties are not designed carefully, it could exacerbate the validator centralization problems already faced by Ethereum today. One possible approach could be to have discriminatory missed slot penalties based on the stake size, assuming the validator cap is raised. This echoes the two-tiered staking and more nuanced staking mechanisms proposed by Vitalik in this post. While such an approach is less egalitarian, professional operators who hold large amounts of ETH and have a highly optimized slot miss rate may be still sufficiently incentivized. This is still dependent on how the missed slot is calculated. There is a possibility that operators may break up their stake across many smaller validators despite the cap raise in an attempt to avoid missed slot penalties.

Concluding thoughts

Through observing the missed slots from various angles, we have evidence to suggest that

1. Slot miss rate will double under competitive relay with single broadcasting, not accounting for timing games.

2. Slot miss rate can be improved exponentially with an increasing number of cooperatively broadcasting relays.

3. Missed slots result in an opportunity cost for the proposers and hence the winning competitive relay.

From these observations, I have proposed two ideas to incentivize relays that enable both competitive and cooperative relaying.

1. Creation of a low-latency relay mesh network for rapid payload distribution for further propagation, aka BaaS.

2. Enshrining missed slot penalty on proposers as a minimal viable enshrinement for relay incentivization, giving more reasons for a proposer to not miss a slot, aka use BaaS.

While relay incentivization is challenging, it's certainly not impossible with coordinated efforts by relay operators and modification to the current incentive for missed slots. I wish this piece uncovers the path to a future with profitable relays and gives some hope to those who are operating or thinking about operating a relay to support Ethereum PBS.

There is much more research work to be done to uncover the path forward for relays & PBS. Here is an incomplete list of some open questions that I have so far:

1. What network topology results in a higher missed slot for single relay broadcasting?

2. Could there be bid biases among builders submitting to multiple relays VS builders that don't? And how does that contribute to the missed slots?

3. How and how much should we penalize proposers for the missed slot without exacerbating validator centralization? Could there be heterogenous penalties applied based on the stake amount, similar to the tiered staking model proposed by Vitalik?

4. Is this relay/acc???

I am looking forward to further discussions among the communities and many work that comes after this. Please also feel free to DM [me](me) for any discussions.

Appendix:

Eq 1

. $R_{Ap} + R_{Yp} = \frac{W_{p}}{W_{\sum}- W_{p}}R_{A}\frac{N}{32} + 512 \frac{W_{p}}{W_{\sum}- W_{p}}R_{Y}$

where $R_{A}\frac{N}{32}$

is the total attestation rewards per slot (proposer reward scale with number of attestors)

Eq 2

. $R_{A}=\frac{14 + 26 + 14}{64} \times 32b$

Eq 3

. $R_{Y} = \frac{W_{y}}{32 \times 512 \times W_{\sum}} Tb$

Eq 4

$$\text{Base reward (b)} = \frac{1,000,000,000 \times 64}{\sqrt{32,000,000,000 \times 900,000}} = 377 \text{ Gwei}$$

Eq 5

$$R_{Ap} + R_{Yp} = \frac{900,000}{32} \times \frac{8}{64- 8} \times \frac{14 + 26 + 14}{64} \times 32b + 512 \frac{8}{64 -8} \times \frac{2}{32 \times 512 \times 64} \times 900,000 \times 32b$$

$$= 40,897,768 + 1,514,732$$

$$= 42,412,500 \text{ Gwei}$$

$$= 0.042 \text{ ETH} \text{ (Using Eq 1, 2, 3, 4 in Appendix)}$$

Relevant resources:

- [[https://ethresear.ch/t/sticking-to-8192-signatures-per-slot-post-ssf-how-and-why/17989/22](https://ethresear.ch/t/sticking-to-8192-signatures-per-slot-post-ssf-how-and-why/17989/22%5C)]

- The Pursuit of Relay Incentivization — Ballsyalchemist

- Timing Games: Implications and Possible Mitigations

- https://www.youtube.com/watch?v=kumD7njaCcU

- Introducing the Validator Gateway: Boost Your Ethereum Validator Rewards - bloXroute

- Increase the MAX_EFFECTIVE_BALANCE – a modest proposal

- [2305.09032] Time is Money: Strategic Timing Games in Proof-of-Stake Protocols

- Sticking to 8192 signatures per slot post-SSF: how and why - #22 by aivarasko

- https://observablehq.com/@libp2p-workspace/performance-dashboard