

# Introduction

Dashboards that investigate and highlight the relationships between different layers of the transaction supply chain is vital in protecting Ethereum's decentralisation and censorship-resistance. [mevboost.pics](#), [mempool.pics](#), [relayscan.io](#), [eigenphi.io](#), and the [Flashbots Transparency Dashboard](#) furnish researchers with critical evidence to scrutinize the activities and dominance of builders and relayers.

However, the searcher sector still lacks illumination. The anonymous nature and variable strategies of searching make searcher identification a challenging task.

Searcher teams that maintain internal searcher datasets have financial incentives to withhold this information. Builders, who may possess searcher datasets (like Titan Builder, who has done [this great public research](#) on searcher dominance), find themselves unable to disclose their datasets fully due to their sensitive position as builders.

[searcherbuilder.pics](#) attempts to collapse the public and private knowledge gap regarding the searcher sector. Specifically, [searcherbuilder.pics](#) answers questions about the state of searcher-builder integration, searcher dominance, and related dynamics. The dashboard differentiates the atomic and non-atomic MEV domains due to their differences in scale and fitting metrics.

In this post, we introduce our methodology behind [searcherbuilder.pics](#) and discuss findings from a recent two-week period. Researchers and other interested parties can then utilize [searcherbuilder.pics](#) and its underlying data to assess the practical risks of vertical integration within Ethereum.

## Summary

[searcherbuilder.pics](#) examines on-chain MEV transaction in both atomic and non-atomic domains. Atomic MEV refers to DEX-DEX arbitrage, sandwiching, and liquidation. Non-atomic MEV refers to CEX-DEX arbitrage. We employ three different metrics to measure the flow from searchers to builders: volume (USD), transaction count, and total bribes (coinbase transfers + priority fees, in ETH).

Notably, we observe that the biggest builders are vertically integrated with the biggest non-atomic searchers: Wintermute is integrated with rsync-builder and Symbolic Capital Partners with beaverbuild.

We also see signs of vertical integration between small builders and small atomic MEV searchers. For example, ~85% of 0xb0bababe's on-chain transactions were captured in blocks produced by boba-builder, when boba-builder is responsible by <1% of all on-chain transactions.

In this post, we explain our methodology behind the dashboard and highlight results from a recent 14-day period (2023-9-30 to 2023-10-12). Unlike [mempool.pics](#), [searcherbuilder.pics](#) solely examines transactions that have landed on-chain and do not rely on mempool or relay bids data. All relevant code can be found in [this repo](#).

## Methodology

### Identifying Atomic MEV Activities & Searcher Addresses

Using [Zeromev's API](#), which employs a slight modification of Flashbots' [mev-inspect-py](#) for MEV detection, we identify atomic MEV transactions in each block. We collect transactions labeled with the `mev_type`

of arb

, frontrun

, backrun

, and liquid

. The smart contract invoked in these transactions, represented by the `address_to`

field returned by the Zeromev API, is the potential MEV searcher address.

From these addresses, we filter out labeled non-MEV smart contracts (such as routers, wash trading bots, telegram bots, etc). Only MEV searchers using proprietary contracts will be detected. Although MEV can be extracted through generic contracts, like Uniswap routers and telegram bots, these opportunities represent an insignificant portion of MEV volume. The set of known contract labels is created by aggregating from multiple sources and active manual inspection.

Zeromev captures a reliable lower bound of active atomic MEV searcher addresses with minimal false positives. Our identification of atomic MEV is ultimately limited by Zeromev's and `mev-inspect-py`'s capabilities, which have [known issues](#)

and miss atomic MEV transactions that fall through their [classification algorithm](#).

## Identifying Non-Atomic MEV Activities & Searcher Addresses

In “[A Tale of Two Arbitrages](#)”, it is estimated that at least “60% of [arbitrage] opportunities (by revenue) are executed via CeFi-DeFi arbitrage”. Capturing such non-atomic MEV activities is the crux of this dashboard.

From all the directional swaps identified by Zeromev, we classify a swap as a CEX-DEX arbitrage if it fulfills one of the following heuristics:

1. It contains an coinbase transfer to the builder (or more generally fee recipient) of the block.
2. It is followed by a separate transaction that is a direct transfer to the builder (a variation of the bribing behaviour above).
3. It is within the [top 10% of the block](#). This aims to capture both CEX-DEX arbitrages that are either bribing solely via gas fees or not bribing at all due to of vertical integration. The heuristic is based on [a demonstrated correlation](#) between top-of-block opportunities and CEX-DEX arbitrage, due to the urgency to extract these MEV opportunities.
4. It interacted with only one protocol. Zeromev has often misclassified some atomic arbitrage as directional swaps; and since atomic arbitrages share the above bribing patterns, they get counted as a CEX-DEX arbitrage. To reduce such false positives, we only look at transactions that are one-hop. While this captures most CEX-DEX arbitrages, those with multi-hops DEX-legs are missed.

We collect the address\_to

field of these transactions and filter out known non-MEV contracts. In the future, we intend to incorporate [price volatility data on leading CEXes](#) to further improve the accuracies of our results.

We also remove any addresses that have been identified as an atomic searcher to further mitigate Zeromev’s occasional misclassification of atomic MEV as swaps. While this means we won’t capture searcher addresses that pursue both atomic & non-atomic MEV opportunities, these addresses are insignificant in number likely due to the need for specialization.

### Note: Not all non-atomic MEV transactions are CEX-DEX arbitrage

Notably, we observed that a very small portion of the non-atomic MEV transactions identified using the above methodology are cross-chain arbitrage rather than CEX-DEX arbitrage.

For example, this Ethereum [transaction](#) that would’ve been picked up by our methodology is actually an arbitrage between the Uniswap pools on Ethereum and Polygon (this is the [Polygon side](#) of the arbitrage). Understanding the size of cross-chain MEV is an interesting open problem space that we may be interested in tackling.

## Metrics for Searcher-Builder Flow

Flow from searchers to builders can be interpreted with three different metrics: volume (USD), transaction count, and total bribe (coinbase transfer + priority fees, in ETH). Each chart can be viewed in each metric using the upper-right toggle.

We recommend volume (USD) as the metric to analyze non-atomic MEV flow. More transactions does not necessarily indicate more dominance for CEX-DEX arbitrages. Given the state of non-atomic searcher-builder integration, we are skeptical that bribe size is correlated with dominance and trade size. Integrated searchers can be over-bribing builders to lend their builder more leverage in the relay auction or under-bribing since their builder can subsidize their bid directly.

In contrast, we recommend transaction count as the best metric for atomic MEV activities. Due to flash loans, volume loses credibility. We don’t recommend total bribes for similar reasons above. Transaction count speaks to the ability for atomic searchers to land on-chain, which is a good proxy for their dominance.

We decided against showing a combined MEV activities. There isn’t a single metric that can satisfactorily represent and compare both MEV domains.