MEV: The First Five Years

James Prestwich

--

Foreword

Five years ago, I wrote a blog post called Miners Aren't Your Friends, introducing MEV to Ethereum dialogue. Back then, we didn't call it MEV. Phil Daian & his co-publishers named "Miner Extractable Value" a year and a bit later in their seminal paper: Flash Boys 2.0. Back then we planned a followup post with python snippets for generating frontrunning transactions on EtherDelta and other mainnet dexes. I got pretty involved in building HTLCs and other cross-chain instruments, and the follow-up post got delayed and then delayed and then back-burnered permanently.

Sitting down to write about MEV today, five years later, is a little surreal. It feels exactly the same as last time. Once again, we're at the start of a long bear market, in the wreckage of another speculative frenzy. The familiar hangover is creeping back in. And just like last time, I'm filled with conviction that next time will be different. By next time we'll have learned from our mistakes and next time we'll build something better. Probably.

Author's notes:

This is not an MEV primer. Many of those have been written, and most are better than I could produce. I assume you're familiar with MEV. If you're not familiar, this is a good time to go read a few blog posts and then come back.

This is a narrative history of MEV. It is not unbiased. It's meandering and a little self-indulgent. I'm a supporting cast member (at best) in this story, but this is my blog so I get to tell it from my perspective.

MEV: The First Five Years

MEV is fundamental to protocol design. Every serious practitioner understands MEV, plans for MEV, and has strong opinions about MEV. From where we are, it's easy to forget that our understanding of MEV is new. The roots are as old as crypto, but the terminology, framing, and tooling are still being invented. The past two years have revolutionized our understanding of MEV. It's easy to forget that MEV started from nothing.

Life before MEV

MEV discourse provenance goes back to early 2010s Bitcoin research on "fee sniping". Fee sniping would later be generalized to the EVM state model and formalized into the "time bandits" attack described in the Flash Boys 2.0 paper. Bitcoin implemented both the first (accidental) consensus-layer MEV mitigation at launch (the 100 block coinbase maturity rule) as well as the first (intentional) client-layer defense against MEV in December 2014 by adding a 1-block timelock to transactions made with the node wallet.

Contention over state and committed-but-unspecified state-transitions produce MEV. Because Bitcoin has basically no shared state over which to contend, and Bitcoin state-transitions are tightly-specified, MEV is usually limited to fee sniping and other sorts of double-spend attempts. Which is to say, Bitcoin miners have very little MEV to extract without attacking the consensus mechanism directly. This makes Bitcoin a uniquely uninteresting chain for MEV research, and we won't spend too much time on Bitcoin or any of its trashy cousins.

The Birth of MEV (2018–2019)

Like any good relationship, MEV needs two things: 1) contention and 2) commitment. First, it needs people vying for control of some public state. MEV needs state that people want, and are willing to pay to get. Second, it needs commitment (prior to execution). MEV needs a period of time where the call to update that contentious, desirable state is committed, but not yet executed. When a user signs and broadcasts a transaction they commit to compete for control of that state. The lag between commitment and execution allows MEV to sneak in on the edges, and touch that state before and afterwards. It pulls on the user's intent, pushes their commitment towards its worst acceptable outcome, and allows the extractor to pocket the spread.

DEXs drip with MEV. They have a perfect recipe for contention: everyone wants to trade on the market, and every trade can compete with every other trade. Self-matched orderbook DEX designs like 2016's EtherDelta allowed miners to get first crack at trades, but no one really did anything with it. A few on-chain auto-matching CLOBs were designed, which seemed crazy at the time, and now feels absolutely bizarre in hindsight. While we knew that eventually gas would cost money, a unit of gas in mid-2018 cost around 1/400th of what it does today and 1/10,000 of what it would last year. We wouldn't earn our visceral understanding of gas price markets until 2020.

The launch of DAI (now SAI) in 2017 introduced liquidations to DeFi. Liquidations introduce massive but infrequent MEV ("spike" MEV). Because an extractor's upside is a percentage of the outstanding debt, outlier CDPs make up an outsize portion of liquidation-related MEV. In the past couple years we've observed oracles cheating to win liquidations on their own protocols and extreme gas auctions to win spike MEV. Spike MEV may also be created by hacks, NFT drops, and other exceptional events. Because it's not divisible into small packets, spike MEV has an outsize effect on protocol operation. Protocol design must account for infrequent, but potentially extremely large, mal-incentives.

DEXs, on the other hand, tend to move with the churn of outside markets creating "flow" MEV from back and forth movements in their markets. Flow MEV is characterized by more frequent, and smaller, MEV packets. AMMs are particularly interesting because their purpose is inseparable from MEV. They exist to track the outside market, and they do so by giving away value to extractors that push the constant function towards "fair" pricing. In effect, they're kept efficient by "arbitrageurs" extracting MEV. The first wave of AMMs (first Bancor, then Uniswap) in mid-to-late 2018 started building flow across their order books. created a corresponding increase in MEV, and drew significant attention to the subject. Of course, back in 2018, Uniswap pools were measured in the single-digit millions (if that!), and there were far fewer assets.

MEV research at this time was nearly non-existent, and of course, we didn't even call it MEV. We mostly called it frontrunning, borrowing the name from TradFi exchanges, "DEX frontrunning" or "liquidation frontrunning." We knew it generalized to more than just order insertion. We understood backrun and sandwich attacks in theory, and we knew that transactions could be delayed or caused to revert if it were profitable for extractors. We had precious little formalism, and no practicum at all.

Ancient DeFi quietly may have built the environment necessary to create MEV. However, no one really bothered. The state wasn't desirable enough yet, so the returns weren't there yet. Because no one had practical experience with this, we didn't even have a good idea of what extraction would look like when it arrived.

Forming Flashbots (2019–2020)

MEV research came into vogue with the publication of Flash Boys 2.0 in April 2019. The paper is excellent. When you have a moment, go re-read it. Then google each of the authors. It is a foundational paper for protocol and mechanism designers. It'll be part of the standard curriculum forever. Without Tina, I think it might have stopped there.

Tina and I met in some VC's office in early 2018. We had been connected by a mutual friend and our pitch meetings were back to back. I was pitching Summa's HTLC variants, which would eventually be canned as we learned the limitations of cross-chain protocols requiring two online parties (and the limitations of market appetite for complexity). Tina was pitching a social blockchain game about milking cows on a farm (no relation to CoW Swap, which came later). We shared an interest in mining derivatives, and had a few ideas on how to approach them. We caught up at every hackathon in the bear market. She has a talent for choosing an idea just ahead of its time, and pulling together a team around it. Her hackathon projects turned into CarboClan which built honeylemon on top of the mining derivatives idea (but without the cows).

Towards the back half of 2019, Tina started organizing events. She put together defi.wtf in just a week and hosted it alongside DevCon Osaka. The next conference, macro.wtf, followed just 3 weeks later. The Pirate Ships started not too long after that. The name "Pi-rate Ship" came from a joke we made at EthDenver in 2019 about flashloans and flashmints. Pirate Ships were salon-style get-togethers hosted by Tina in San Francisco, New York, and elsewhere. You could drop by the ship and find an eclectic collection of researchers, engineers, operators, grifters, and other crypto folks. Some would have researchers-in-residence, some would have a topic or theme, some would just be a week in Puerto Rico with whoever happened to be there. Ships ran from late 2019 throughout the pandemic, pulling in the most persistent nomads.

Flashbots owes its existence to Pirate Ships far more than it does to the Flash Boys 2.0 paper. It was born as a perpetual salon. The MEV Ship was formed in mid 2020, and expanded naturally from an in-person space to a digital collective. The other Flashbots founders and stewards — Stephen, Phil, Alex — were pulled into the Ship (you can still find a few allusions to this in the Flashbots docs). Eventually the Ship formalized, changed their name to Flashbots, put on a cute robot emoji, and slipped into immortality.

That entire year was one long breakout moment. Flashbots Research (née the MEV Ship) debuted and propelled MEV to the center of the Ethereum dialogue. DeFi Summer had started, but we didn't know what it was yet. The release of mev-explore and mev-inspect put rough numbers on the MEV available on-chain, and all of a sudden it was tangible. No longer a theory. With the announcement of Flashbots Auction that fall, it became clear that MEV extraction tooling was just around the corner. There was money on the ground and a race to pick it up.

The MEV Ship provided the foundation of Flashbots and set the culture of collaborative research and experimentation that drives the organization. The clear mission derived from that culture has defined the MEV narrative for years. Productized MEV extraction was inevitable, but the Flashbots culture and ethos were not.

Professionalization (2021 — present)

The halo around the MEV ship and Flashbots attracted (largely non-traveling) technical talent, of course. The Flashbots Auction (mev-geth and the flashbots relay) launched in January 2021 — barely 3 months after it was announced. The timing could not have been more perfect. Mev-geth launched right at the start of the biggest runup in Ethereum's history and (naturally) the correspondingly large increase in MEV generation.

Before Flashbots Auction, a few brave Searchers extracted MEV by broadcasting transactions to the txpool with specific gas prices. It was unreliable and inefficient. It required the searcher to see a transaction in the txpool, simulate, extract, and broadcast within the span of a block. Pre-1559 fee semantics and the general unreliability of the txpool complicated the process. Frontrunning transactions would select fees to try to execute immediately before the target, assuming blocks were fee sorted. Spike MEV transactions would simply pay egregious gas prices. Because the searchers used the public txpool, there were fierce real-time auctions. A mess, but a highly entertaining one.

Flashbots Auction's relative simplicity drove its adoption. It boils down to a single RPC endpoint that allows searchers to submit bundles. Bundles ought to be included in the block exactly as specified, and reverts should cause the bundle to be dropped. This creates a neat interface for specialization. Searchers hunt MEV, trap it, and secure it in nice bundles, and pass it off to miners for inclusion in the chain. The Flashbots Auction quickly came to define extraction, to the point where other systems are barely relevant.

Later extensions of the Flashbots Auction, as we see in mev-boost, separate the "Builder" from the "Producer," (a generalized term for miners and stakers). The first time I heard about proposer-builder separation ("PBS") was EthBerlin 2019. Will, then working on eth2, taught me a lot about designs for state witness production and updating in stateless Ethereum designs. I can't remember what we called it back then, but it was conceptualized as a protocol-mandated specialized node with extra responsibilities. Stateless Ethereum died of Covid. MEV PBS — as instantiated in mev-boost — lives outside the protocol. A market-based specialization, rather than one specified by the protocol.

Today, we have a relatively mature MEV supply chain. Searchers mine transaction flow for MEV, competing with each other for the spikes and flow. Searching requires skill and extreme specialization. They are secretive because they need to keep their edge over other searchers. Builders accept bundles from searchers and build them into blocks. Searchers and Builders have a symbiotic relationship. Builders rely on Searchers to extract, and Searchers rely on Builders to honestly include the bundle contents.

Builders purchase the right to include their block in the main chain from a Proposer (a generalized term for miner or staker). Proposers have the protocol-granted right to add blocks to the chain, and select the builder block which pays the most MEV. Users create MEV, Searchers extract it. Searchers pay Builders; Builders pay Proposers. The gears are greased, the blocks are assembled, the transactions go into the chain. Forever and ever, Amen.

While we're on the topic, a quick aside: Bundles have an unintended side-effect: they cause Searchers to incidentally subsidize transaction fees for the extraction target. When a Searcher includes a user transaction in a bundle, that user transaction confirms earlier and more reliably than it otherwise would. The MEV extracted by the Searcher is partially paid down the supply chain to the Builder and Proposer. This effectively converts some of the extracted MEV into transaction fees for the bundle. In essence, extracted MEV pays a "shadow fee" to the Proposer. The priority of a transaction is determined by its MEV first, and its protocol fee second.

Extraction accidentally undermined EIP-1559. Non-bundle transactions use a per-gas tip on top of the basefee. Bundles, on the other hand, are still effectively a first-price auction with flat tips on the basefee, not per-gas tips. Which is to say, there are now two co-dependent fee markets. One shadow market for MEV transactions, and a regular market for everything else. MEV buys priority. We (confidently and correctly) predicted these shadow fee markets well in advance of 1559 (and of course, every other fee mechanism). What we didn't anticipate was using MEV market rails. Because extraction runs behind the curtain, users benefit from the side market without even being aware of its existence.

MEV-driven PBS is likely not, from a mechanism-design perspective, "secure" or "incentive-compatible." The Flashbots MO so far has been to rely on honesty assumptions to fill in gaps in incentive mechanisms. They then establish those assumptions as market-norm behavior. Why don't Builders break up bundles and take the MEV? Why don't Builders include reverting bundles? Because violating those market norms would cause Searchers to stop using the relay. We bet that the iterated game will be more profitable than a game-ending grab at bundled MEV. I am not sure that it will hold up to spike MEV and vertical integration in the long run.

Incentive incompatibility of the auction can safely be ignored for a while. Market norms suffice. Mechanism design bows to the pressure of the markets. Margin compression is the only relevant problem for the MEV supply chain today. Because a Proposer has the exclusive right to select a block, Builders must compete on price. They are forced to give up more and more of their share of MEV to the Proposers, and take more from the Searchers to pay for it. MEV margins are already collapsing, with Proposers taking the lion's share.

Fundamentally, Proposers extract rent. Builders and Searchers don't have the option to purchase from anyone else, and cannot refuse to buy. Proposers have been granted an exclusive right by the protocol to select the next block, and can exercise that right without oversight. To paraphrase a rather famous politician: "Proposers got this thing and it's fucking golden. They're not just giving it up for fucking nothing." Rod, like the Proposer, was granted something incredibly valuable by an arcane protocol quirk. This isn't a moral or natural right. It's granted by the rules of the protocol, which can be changed

by humans. Regardless, Proposers are paid for extending the chain, and then Builders pay the Proposers to extend it in a specific way. Awful nice to get paid twice.

Before MEV extraction, ordering was a perfunctory role; the inflationary block subsidy was the main incentive. As such, the right was assigned willy-nilly, for free. That's why Proposers pay nothing for this right. It used to be worth next to nothing. Now, the ordering right includes the rights to a vast amount of MEV, and this asymmetric power relationship warps the MEV supply chain. Research in this area is ongoing. It seems likely we'll see in-protocol mechanisms proposed to address this.

Sunny advocates for threshold encryption for in-flight transactions. Encrypted transactions prevent Proposers from learning the MEV value of transactions before they are ordered. Personally, I don't believe in MEV expurgation. I'm interested in designs that include MEV in the fork-choice rule, selling the ordering right to the highest bidder and keeping the bulk of the extracted value for in-protocol re-allocation. Embrace, extract, expropriate.

MEV: The First Five Years

So where does that leave us? Five years ago, MEV was a curiosity. It was a researcher dinner conversation, found at the bottom of the wine bottle, after all the real business had been done. MEV waited for the right combination of DeFi activity and practical experience. Flashbots captured MEV research mindshare just as we entered the 2021 bull market, and rode the wave all the way up.

Now, we hold MEV symposia. We get Searchers and devs together to talk about MEV-aware protocol design. The market structure has settled on a 3-role configuration, and is starting to test its limits. While there's still respect for the Pirate Ships' ideals, extraction on Ethereum has professionalized. Going forward, MEV won't be defined by the ethos or the research. MEV's story belongs to profit margins now.

Throughout its history, people (including me) have called MEV dangerous or evil. The word "theft" has been bandied about, more than a few times. These days, I think it's unreasonable to make a value judgment. MEV is a fact. No amount of philosophizing, research, or rube goldberging will change that. This is why I'm happy to see a competitive MEV extraction supply chain. Professionalized MEV is predictable MEV; predictable MEV is useful MEV. We aren't smart enough to plan a market in advance, so let's take utility where we find it.

Acknowledgements

Nobody. I'll see you all in hell.