

Layer 2 on Hyperliquid: A Cohesive Analysis of Opportunities and Limitations

Souren Khetcho
Independent Research Efforts
`souren@avo.so`

June 2025

1 Introduction

Introducing a Layer 2 (L2) solution on Hyperliquid presents significant potential for building advanced trading applications that leverage Hyperliquid’s impressive performance characteristics (200,000 TPS, sub-second finality). Rather than replacing the existing strengths of Hyperliquid, an L2 would amplify them by enhancing scalability, enabling specialized execution environments, and facilitating broader ecosystem integration. This paper explores the opportunities and limitations by evaluating the technical necessity and feasibility of an L2 infrastructure for Hyperliquid, particularly in the context of advanced AI-driven trading applications.

2 Hyperliquid’s Foundation for L2 Innovation

Hyperliquid has established itself as the dominant on-chain derivatives platform, capturing 70% market share with its custom HyperBFT consensus and dual architecture combining HyperCore (native trading) with HyperEVM (smart contracts). However, even with its impressive 200,000 TPS and sub-second finality, **Hyperliquid faces scaling constraints** as it approaches Tendermint limits, geographic latency issues with Tokyo-concentrated validators, and centralization concerns with only 16 validators. An L2 could address these limitations while unlocking entirely new application categories.

3 Revolutionary Trading Applications Enabled by L2

3.1 Institutional-Grade High-Frequency Trading Infrastructure

While Hyperliquid already outperforms most platforms with 200k TPS, **an L2 could enable true millisecond latency trading** that rivals centralized exchanges. This would support sophisticated applications like real-time cross-market arbitrage engines that simultaneously monitor Hyperliquid’s native orderbooks, external DEXs, and centralized exchanges to execute complex arbitrage strategies within microseconds. Traditional HFT firms could deploy algorithmic trading strategies previously impossible in DeFi, including statistical arbitrage, momentum trading, and market-neutral strategies that require sub-100ms execution times. The L2 could host professional market-making infrastructure where institutional participants run sophisticated quoting engines that dynamically adjust spreads based on volatility, inventory, and market conditions. Unlike current AMM-based market making, this would enable order book-native strategies that provide deeper liquidity and tighter spreads, particularly valuable given Hyperliquid’s dominance in derivatives trading.

3.2 Advanced Derivatives and Structured Products Ecosystem

Building on Hyperliquid’s perpetuals leadership, **an L2 could enable complex derivatives** that require intensive computation currently prohibitive even at Hyperliquid’s performance levels. This includes **multi-underlying options strategies** where traders can create complex positions across multiple assets with real-time Greeks calculation and dynamic hedging. **Structured products** like barrier options, variance swaps, and volatility derivatives could execute continuously with real-time risk management. **Cross-protocol derivatives** represent a unique opportunity - instruments that derive value from positions across Hyperliquid’s native orderbooks and external DeFi protocols. For example, **basis trading products** that automatically capture funding rate differences between Hyperliquid perpetuals and other platforms, or **volatility products** that trade the spread between implied volatility on Hyperliquid and realized volatility in spot markets.

3.3 Micro-Transaction and Social Trading Applications

Micro-betting on price movements could enable users to place cents-level wagers on short-term price direction, creating engaging prediction markets around trading outcomes. **Creator monetization tools** could leverage Hyperliquid’s trading data to enable performance-based subscriptions where traders pay content creators based on realized alpha generation. **Real-time tip flows** could reward successful trade calls instantly, creating new economic models for trading education and signals.

4 Enhanced Functionality for Existing Hyperliquid Features

4.1 Intelligent Order Management and Execution

An L2 could dramatically enhance Hyperliquid’s native trading with **AI-powered order execution** engines that optimize trade execution across time. These systems could analyze market microstructure, predict optimal execution timing, and automatically split large orders to minimize market impact - functionality requiring continuous computation that would benefit from L2’s additional throughput capacity.

Advanced conditional orders could execute complex multi-step strategies automatically. For example, portfolio rebalancing engines that monitor positions across multiple Hyperliquid markets and automatically adjust allocations based on correlation changes, volatility shifts, or technical indicators, all executing at speeds impossible without dedicated L2 infrastructure.

4.2 Enhanced Liquidity and Market Depth

Aggregated liquidity pools on the L2 could combine liquidity from Hyperliquid’s native orderbooks with external sources, creating deeper markets particularly valuable for institutional-size trades. This could include cross-chain liquidity aggregation where the L2 seamlessly accesses liquidity from Arbitrum, Optimism, and other networks (possibly non-EVM chains such as Solana, although the technical feasibility has not been assessed yet) while maintaining Hyperliquid’s performance characteristics.

Dynamic market making strategies could adjust liquidity provision in real-time based on market conditions, trader behavior, and risk parameters. Unlike static AMM curves, these could provide adaptive spreads that tighten during normal conditions and widen during volatility, optimizing capital efficiency for liquidity providers.

5 Reality Check: L2 Necessity for Hyperliquid AI Trading

Hyperliquid’s high-performance architecture demonstrates that **Layer 2 solutions are not inherently necessary** for AI trading applications, but specific technical requirements may justify specialized infrastructure in certain scenarios. This analysis reveals that the decision hinges on precise computational demands rather than theoretical scaling needs.

5.1 Critical performance gap between requirements and reality

The fundamental mismatch lies in the latency requirements versus the capabilities of the blockchain. **High-frequency AI trading demands submillisecond response times** (1-10 microseconds for competitive HFT strategies), while even Hyperliquid’s optimized architecture achieves 200-900 millisecond latencies. This represents a 1000x performance gap that no L2 solution can bridge through consensus optimization alone. Research shows that top HFT firms require actionable latency of 13.9 nanoseconds (achieved with specialized FPGAs) to 50-80 microseconds (GPU-based inference). Hyperliquid’s current theoretical 200,000 TPS throughput with sub-second finality, while impressive for blockchain standards, remains fundamentally unsuitable for latency-critical AI trading strategies that dominate 80% of modern high-frequency trading volume. However, it’s important to recognize that HFT represents only a niche segment of the broader trading landscape, and optimizing for ultra-low latency should be treated as a secondary priority rather than a core design goal.

5.2 Hyperliquid’s architecture strengths and limitations

Hyperliquid’s technical capabilities present a **mixed picture for AI trading applications**. The platform achieves genuine high performance with 20,000 TPS operational capacity (not the theoretical 200k claimed), supporting \$1-10 billion daily trading volumes through its custom HyperBFT consensus mechanism. The dual architecture approach with HyperCore for native trading and HyperEVM for smart contracts provides flexibility for different computational needs. However, **severe centralization constraints** undermine the platform’s decentralized security model. With 16 validators and 81% foundation control, Hyperliquid operates more like a high-performance centralized system than a decentralized blockchain. This centralization enables the current performance but creates single points of failure and regulatory risks that may not justify the complexity of blockchain infrastructure for AI trading applications. The HyperEVM smart contract layer can theoretically handle many proposed L2 use cases but **lacks specialized AI optimizations** such as custom opcodes, GPU integration, or optimized data structures for machine learning workloads. Standard EVM limitations apply to complex AI computations, making it unsuitable for on-chain AI inference without substantial modifications.

5.3 AI trading requires hybrid architectures, not L2 solutions

Analysis of successful AI trading implementations reveals that **hybrid architectures consistently outperform blockchain-native approaches**. Projects like Numerai execute AI computation entirely off-chain while using blockchain only for token incentives and result verification. This pattern emerges because AI trading demands **specialized computational environments** that general-purpose blockchains cannot efficiently provide.

The computational requirements are substantial: GPT-3 level models require **350 TFLOPs per inference query**, while competitive trading strategies need **memory bandwidth of 1,555 GB/s** (achievable with specialized GPUs but not blockchain VMs). These requirements suggest that application-specific infrastructure outside blockchain constraints may be more appropriate than L2 solutions that attempt to optimize blockchain execution.

Real-world AI trading systems achieve competitive performance through FPGA acceleration, specialized neural processing units, and custom silicon optimized for specific algorithms. L2 solutions cannot replicate

this level of hardware specialization, suggesting that the bottleneck is architectural rather than consensus-related.

5.4 L2 solutions create more problems than they solve

The research reveals that **L2 implementations introduce significant technical overhead** without addressing core AI trading requirements. ZK-rollups require computationally intensive proof generation that adds latency rather than reducing it. Optimistic rollups introduce 7-day withdrawal delays and complex fraud proof mechanisms that conflict with AI trading’s need for immediate settlement.

Bridge security represents a critical vulnerability, with 69% of DeFi exploits in 2022 targeting cross-chain bridges representing over \$2 billion in losses. For AI trading applications handling significant capital, these security risks outweigh potential performance benefits. The additional **trust assumptions** required for L2 infrastructure conflict with the security requirements of trading applications.

Furthermore, L2 solutions create **fragmentation of composability** that limits AI agents’ ability to interact with multiple liquidity sources simultaneously. Many trading agents would often require cross-venue arbitrage and complex multi-asset strategies that become more difficult when isolated on specialized L2 infrastructure.

5.5 When specialized infrastructure is justified

The analysis identifies specific scenarios where **application-specific architecture** (not necessarily L2) becomes technically necessary. **Ultra-high frequency trading** requiring submillisecond execution cannot operate on any blockchain architecture and requires dedicated infrastructure similar to traditional financial markets.

Complex AI model deployment with specialized hardware requirements (GPU clusters, custom ASICs) may benefit from **dedicated execution environments** rather than general-purpose blockchain VMs. However, this points toward **off-chain computation with on-chain verification** rather than L2 blockchain solutions.

Regulatory compliance for AI model execution and data privacy may require **specialized environments** with audit trails and controlled access. These requirements suggest **permissioned computing environments** rather than public L2 blockchains.

5.6 Technical evidence supports L1 optimization over L2 deployment

Comparative analysis of successful implementations demonstrates that **L1 optimization often provides superior results** to L2 deployment. Projects like Jupiter on Solana achieve competitive performance by leveraging Solana’s native capabilities rather than adding L2 complexity. Uniswap’s continued success on Ethereum L1 despite L2 availability shows that network effects and composability often outweigh raw performance metrics.

The performance benchmarking reveals that **most blockchains operate far below theoretical capacity** - Solana achieves 1,053 TPS average versus 65,000 TPS theoretical maximum. This suggests that L1 optimization has significant headroom before L2 solutions become necessary.

Cost analysis shows that Solana’s inexpensive transactions already achieve the cost efficiency that many L2 solutions target, while maintaining stronger security guarantees and avoiding bridge risks.

5.7 Consideration: Build DeFi applications, not L2 infrastructure

The technical evidence strongly supports **building sophisticated DeFi applications on Hyperliquid’s existing infrastructure** rather than developing L2 solutions. Hyperliquid’s 20,000 TPS capacity with

sub-second finality already exceeds the requirements of most AI trading applications that don't require sub-millisecond execution.

For AI trading applications: Use Hyperliquid's HyperEVM for coordination and settlement while performing AI computation off-chain with specialized hardware or LLMs. This hybrid approach captures blockchain benefits (transparency, settlement finality) while avoiding blockchain limitations (computational constraints, latency overhead).

For high-frequency trading: Recognize that blockchain architecture fundamentally cannot compete with traditional financial infrastructure for latency-critical strategies. Focus on blockchain-native use cases like decentralized market making, cross-chain arbitrage, and transparent algorithmic trading rather than attempting to replicate traditional HFT on blockchain infrastructure.

The analysis reveals that L2 solutions for AI trading represent premature optimization that adds complexity without addressing fundamental architectural limitations. The more effective approach focuses on leveraging blockchain strengths (decentralization, transparency, programmability) while using appropriate off-chain infrastructure for computationally intensive AI operations.