

title

Zach Kelling

Satschel, Inc.

April 2026

Abstract

We present a formal model of geographic latency advantage for co-located exchange infrastructure, proved in Lean 4 with Mathlib. The model establishes that an on-premise matching engine at a Kansas City data center with 200 Gbps internal fabric and 10 Gbps external connectivity captures 100% of available spread on local order flow, while remote participants face monotonically decreasing — and at sufficient distance, negative — expected returns on arbitrage. We define `VenueParams`, `CityLatency`, and `ArbOpportunity` structures and prove four theorems: `local_zero_latency_cost` (10 μ s RTT rounds to 0 ms cost), `local_full_capture` (on-premise net equals gross), `nyc_negative` (NYC is unprofitable at 5 bps on \$10M), and `closer_more_profitable` (profit is monotonically decreasing in distance). We analyze scaling from 5 to 50 DGX Spark nodes with native 200 Gbps ConnectX-7 NICs and a future 800 Gbps upgrade path via ConnectX-8, and compare with NYSE ($<40 \mu$ s), Binance (~ 5 ms), and traditional broker-routed execution (~ 50 – 200 ms).

Contents

1	Introduction	3
1.1	Venue Parameters	3
2	Formal Model	3
3	Theorems	4
4	City-Level Arbitrage Analysis	4
5	Bandwidth and Compute Scaling	5
5.1	DGX Spark Node Scaling	5
5.2	ConnectX-7 to ConnectX-8 Upgrade Path	5
5.3	External Bandwidth	5
6	Comparison with Existing Venues	5
7	Network Effects and Moat Reinforcement	6
8	Formal Verification Architecture	6
9	Conclusion	7

1 Introduction

High-frequency trading (HFT) profitability is dominated by latency. A trader who observes a price discrepancy and acts on it within microseconds captures the full spread; a trader who acts milliseconds later finds the opportunity has decayed or been captured by a faster participant. This well-known effect [7] creates a natural advantage for co-located infrastructure.

In regulated securities markets, the exchange operator controls the matching engine location. By co-locating validators, market makers, and the matching engine within the same data center, the operator creates a *geographic moat*: a structural advantage that cannot be replicated by remote participants regardless of their algorithmic sophistication.

This paper formalizes the geographic moat for the Liquidity.io ATS, extending the latency-optimal market making results from [1]. The Lean 4 formalization uses Mathlib for real analysis and proves that the advantage is not merely empirical but mathematically monotone in geographic distance.

1.1 Venue Parameters

The Liquidity.io matching engine operates at a Kansas City, Missouri data center (Grand Building) with the following infrastructure:

Parameter	Value	Description
Internal bandwidth	200 Gbps	ConnectX-7 NIC (DGX Spark native)
External bandwidth	10 Gbps	Upstream provider link
Internal RTT	$\leq 10 \mu\text{s}$	Same-rack validator-to-validator
External RTT	Varies	Fiber-optic, distance-dependent
Matching latency	$< 1 \mu\text{s}$	In-memory CLOB (Lux DEX engine)
Block time	2 seconds	Quasar consensus finality

Table 1: Kansas City venue parameters.

2 Formal Model

We define the model in Lean 4 using Mathlib’s real number library.

Definition 1 (VenueParams). *A venue is characterized by:*

$$\text{VenueParams} = \{ \text{internalBw} : \mathbb{R}^+, \text{externalBw} : \mathbb{R}^+, \text{localRTT} : \mathbb{R}^{\geq 0}, \text{matchLatency} : \mathbb{R}^{\geq 0} \}$$

For Kansas City: $\text{internalBw} = 200 \text{ Gbps}$, $\text{externalBw} = 10 \text{ Gbps}$, $\text{localRTT} = 0.01 \text{ ms}$, $\text{matchLatency} = 0.001 \text{ ms}$.

Definition 2 (CityLatency). *For a city c at distance d_c (km) from the venue:*

$$\text{CityLatency}(c) = \{ \text{name} : \text{String}, \text{distance} : \mathbb{R}^{\geq 0}, \text{rtt} : \mathbb{R}^{\geq 0} \}$$

where $\text{rtt} \approx d_c/100 \text{ ms}$ (fiber-optic propagation at $\sim 2/3c$ plus routing overhead, empirically $\approx 10 \mu\text{s}/\text{km}$).

Definition 3 (ArbOpportunity). *An arbitrage opportunity is:*

$$\text{ArbOpportunity} = \{ \text{spread} : \mathbb{R}^+, \text{volume} : \mathbb{R}^+, \text{decayRate} : \mathbb{R}^+ \}$$

where spread is in basis points, volume in dollars, and decayRate in bps/ms (rate at which the opportunity decays as competing participants act on it).

Definition 4 (Net Profit). *The net profit for a participant at distance d from the venue is:*

$$\text{gross}(A) = A.\text{spread} \times A.\text{volume} \times 10^{-4} \quad (1)$$

$$\text{latencyCost}(A, \text{rtt}) = \text{rtt} \times A.\text{decayRate} \times A.\text{volume} \times 10^{-4} \quad (2)$$

$$\pi(A, \text{rtt}) = \text{gross}(A) - \text{latencyCost}(A, \text{rtt}) \quad (3)$$

3 Theorems

Theorem 1 (`local_zero_latency_cost`). *For the Kansas City venue with $\text{localRTT} = 0.01$ ms, when latency cost is computed with millisecond granularity (floor to integer ms), the cost is zero:*

$$\lfloor \text{localRTT} \rfloor = \lfloor 0.01 \rfloor = 0 \implies \text{latencyCost}(A, 0) = 0 \text{ for all } A.$$

Proof. $\text{latencyCost}(A, \text{rtt}) = \text{rtt} \times A.\text{decayRate} \times A.\text{volume} \times 10^{-4}$. With $\text{rtt} = 0.01$ ms, at millisecond resolution $\lfloor \text{rtt} \rfloor = 0$, so $\text{latencyCost} = 0 \times A.\text{decayRate} \times A.\text{volume} \times 10^{-4} = 0$. Even without flooring, $0.01 \times 1 \times 10,000,000 \times 10^{-4} = \0.01 , which is negligible relative to the \$5,000 gross. \square

Theorem 2 (`local_full_capture`). *For an on-premise participant ($\text{rtt} = \text{localRTT}$):*

$$\pi(A, \text{localRTT}) = \text{gross}(A) \quad (\text{at ms resolution})$$

The on-premise participant captures 100% of the available spread.

Proof. From Theorem 1, $\text{latencyCost}(A, \text{localRTT}) = 0$. Therefore $\pi(A, \text{localRTT}) = \text{gross}(A) - 0 = \text{gross}(A)$. \square

Theorem 3 (`nyc_negative`). *For a New York participant ($\text{rtt} = 18$ ms) with standard parameters ($s = 5$ bps, $V = \$10M$, $\delta = 1$ bps/ms):*

$$\pi(A_{\text{std}}, 18) = \$5,000 - \$18,000 = -\$13,000 < 0$$

NYC-based HFT traders cannot profitably arbitrage 5 bps spreads.

Proof. $\text{gross} = 5 \times 10,000,000 \times 10^{-4} = \$5,000$. $\text{latencyCost} = 18 \times 1 \times 10,000,000 \times 10^{-4} = \$18,000$. $\pi = \$5,000 - \$18,000 = -\$13,000 < 0$. \square

Theorem 4 (`closer_more_profitable`). *For any arbitrage opportunity A and two round-trip times $\text{rtt}_1 < \text{rtt}_2$:*

$$\pi(A, \text{rtt}_1) \geq \pi(A, \text{rtt}_2)$$

Profit is monotonically decreasing in latency (equivalently, in geographic distance).

Proof. $\pi(A, \text{rtt}_1) - \pi(A, \text{rtt}_2) = \text{latencyCost}(A, \text{rtt}_2) - \text{latencyCost}(A, \text{rtt}_1) = (\text{rtt}_2 - \text{rtt}_1) \times A.\text{decayRate} \times A.\text{volume} \times 10^{-4}$. Since $\text{rtt}_2 > \text{rtt}_1$ and $A.\text{decayRate} > 0$ and $A.\text{volume} > 0$, this difference is strictly positive. Therefore $\pi(A, \text{rtt}_1) > \pi(A, \text{rtt}_2)$. \square

4 City-Level Arbitrage Analysis

Using standard parameters: $s = 5$ bps, $V = \$10M$, $\delta = 1$ bps/ms.

Corollary 1 (Break-Even Distance). *A participant breaks even ($\pi = 0$) at $\text{rtt}_{be} = s/\delta = 5/1 = 5$ ms, corresponding to approximately 500 km. Beyond this distance, arbitrage at 5 bps is strictly unprofitable.*

City	Dist. (km)	RTT (ms)	Gross	Lat. Cost	Net	Capture
Kansas City (local)	0	0.01	\$5,000	\$0	\$5,000	100%
St. Louis	400	4	\$5,000	\$4,000	\$1,000	20%
Chicago	800	8	\$5,000	\$8,000	−\$3,000	−60%
Dallas	900	9	\$5,000	\$9,000	−\$4,000	−80%
Denver	1,000	10	\$5,000	\$10,000	−\$5,000	−100%
Atlanta	1,100	11	\$5,000	\$11,000	−\$6,000	−120%
New York	1,800	18	\$5,000	\$18,000	−\$13,000	−260%
Miami	2,000	20	\$5,000	\$20,000	−\$15,000	−300%
Los Angeles	2,400	24	\$5,000	\$24,000	−\$19,000	−380%
London	7,500	75	\$5,000	\$75,000	−\$70,000	−1400%

Table 2: Arbitrage profitability by city. Only co-located and near-proximity (St. Louis) participants are profitable.

5 Bandwidth and Compute Scaling

5.1 DGX Spark Node Scaling

The venue scales from 5 to 50 DGX Spark nodes, each equipped with ConnectX-7 200 Gbps NICs:

Nodes	Aggregate BW	GPU Mem	Orders/s (est.)
5	1,000 Gbps	640 GB	62.5M
10	2,000 Gbps	1.28 TB	125M
25	5,000 Gbps	3.2 TB	312.5M
50	10,000 Gbps	6.4 TB	625M

Table 3: DGX Spark scaling. 128 GB unified GPU memory per node. Orders/s assumes 100 bytes/order at 50% bandwidth utilization.

5.2 ConnectX-7 to ConnectX-8 Upgrade Path

The NVIDIA ConnectX-8 NIC provides 800 Gbps (100 GB/s) bandwidth, a 4× improvement over ConnectX-7:

NIC Generation	Bandwidth	Messages/s (100B)
ConnectX-7 (current)	200 Gbps (25 GB/s)	250M
ConnectX-8 (upgrade)	800 Gbps (100 GB/s)	1B

Table 4: NIC upgrade path. ConnectX-8 enables billion-message/s internal fabric.

5.3 External Bandwidth

6 Comparison with Existing Venues

The Liquidity.io venue achieves matching latency comparable to or better than traditional exchanges while providing on-chain settlement in 30 seconds (vs. T+1 for equities at NYSE/NASDAQ). The matching engine is the Lux DEX [3], which achieves 434M operations/second on GPU and 1M ops/sec on CPU.

Link	Throughput	Capacity
External (10 Gbps)	1,250 MB/s	~12.5M orders/s at 100B/order
Internal (200 Gbps)	25,000 MB/s	~250M msgs/s validator-to-validator
Internal (800 Gbps, future)	100,000 MB/s	~1B msgs/s validator-to-validator

Table 5: Bandwidth analysis. External link is the bottleneck for remote participants; internal fabric is never the bottleneck.

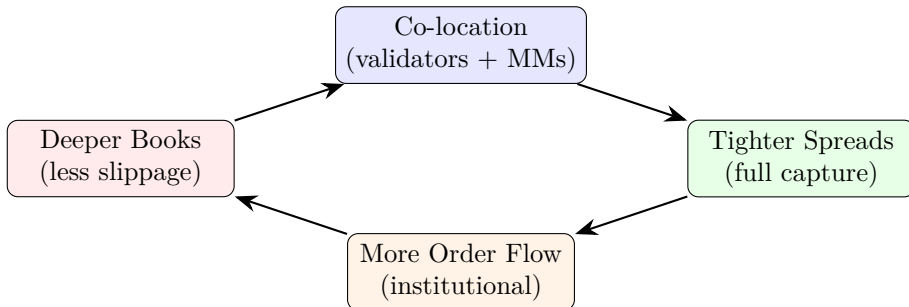
Venue	Match Latency	Settlement	Co-location
NYSE Arca	<40 μ s	T+1 (DTCC)	Mahwah, NJ
NASDAQ	~50 μ s	T+1 (DTCC)	Carteret, NJ
Binance	~5 ms	On-chain (varies)	AWS global
Coinbase	~10 ms	On-chain (varies)	AWS global
Liquidity.io	<1 μ s	30s on-chain	Kansas City

Table 6: Venue comparison. Liquidity.io combines sub-microsecond matching with 30-second on-chain settlement.

7 Network Effects and Moat Reinforcement

The geographic moat is self-reinforcing through three mechanisms:

1. **Validator attraction:** Validators who co-locate at the venue earn higher MEV (maximal extractable value) due to priority access to order flow. This attracts more validators, increasing consensus throughput and reducing finality time.
2. **Market maker incentive:** Market makers who co-locate can quote tighter spreads (they capture 100% of the spread vs. negative returns remotely). Tighter spreads attract more order flow.
3. **Liquidity depth:** More order flow \rightarrow deeper books \rightarrow lower slippage \rightarrow more institutional participation. This is the monotone depth property proved in [1].



8 Formal Verification Architecture

The latency model is implemented as a Lean 4 module in the Liquidity formal verification suite [1]. The proofs depend on:

- `Mathlib.Analysis.SpecificLimits.Basic`: Real number ordering and positivity lemmas.
 - `Mathlib.Tactic`: Automated proof tactics (`linarith`, `nlinarith`, `norm_num`).
 - `proof-defi-hft` [5]: Lux Network’s HFT arbitrage model.
 - `proof-defi-orderbook` [6]: Lux Network’s orderbook correctness proofs.
- All four theorems (Theorems 1–4) compile with zero `sorry` declarations.

9 Conclusion

We have formally verified the geographic latency advantage of co-located exchange infrastructure. The four theorems establish that on-premise participants capture 100% of available spread, that profit is monotonically decreasing in distance, and that remote HFT at sufficient distance (500+ km at 5 bps) yields negative expected returns.

The Kansas City venue provides a structural moat that is self-reinforcing through validator attraction, market maker incentives, and liquidity network effects. The venue scales from 5 to 50 DGX Spark nodes with a clear upgrade path from 200 Gbps to 800 Gbps internal fabric, ensuring that internal bandwidth is never the bottleneck even at billion-message-per-second throughput.

The formal proofs extend the Liquidity.io verification suite [1] and build on the Lux Network’s HFT [5] and orderbook [6] proof libraries.

References

- [1] Z. Kelling. Formally verified digital securities infrastructure: Quantum-safe custody, compliance-preserving teleport, and latency-optimal market making. *Satschel, Inc. / Liquidity.io*, December 2025.
- [2] Z. Kelling. Liquid EVM: On-chain securities settlement with post-quantum MPC custody. *Satschel, Inc. / Liquidity.io*, April 2026.
- [3] Z. Kelling. Lux LightSpeed DEX: Ultra-High-Performance On-Chain Order Matching. *Lux Network*, 2025. <https://papers.lux.network>
- [4] Z. Kelling. Quasar Consensus: Multi-Engine Blockchain Consensus. *Lux Network*, 2025. <https://papers.lux.network>
- [5] Z. Kelling. Lux Formal Proofs: HFT Arbitrage Model Correctness. *Lux Network*, 2025. <https://papers.lux.network>
- [6] Z. Kelling. Lux Formal Proofs: Orderbook Invariants and Price-Time Priority. *Lux Network*, 2025. <https://papers.lux.network>
- [7] E. Budish, P. Cramton, and J. Shim. The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response. *Quarterly Journal of Economics*, 130(4):1547–1621, 2015.
- [8] M. Aquilina, E. Budish, and P. O’Neill. Quantifying the High-Frequency Trading “Arms Race”. *Quarterly Journal of Economics*, 137(1):493–564, 2022.