NeuralMesh™ Scales High Frequency Trading Execution

NeuralMesh™ delivers deterministic microsecond-scale latency high-frequency traders need to execute with predictable performance and capture market moves first.

Challenges

- Latency variance kills execution
- Legacy NAS relies on probabilistic throughput
- GPU resources are underutilized when latency spikes during peak windows
- Petabyte data strains capacity
- Multi-exchange scaling adds variance

Solution

NeuralMesh™ is a fully distributed, shared-nothing architecture with architectural intent for deterministic performance. Every server handles data, metadata, and I/O directly with no bottleneck nodes, eliminating latency variance during concurrent load built for always-on trading workloads.

NeuralMesh Axon™ runs directly on the NVMe storage already inside your GPU servers, moving market data from ingestion to order execution with no copies.

Benefits

- Optimize trading during peak times
- Execute trades with predictable latency
- Process data deterministically
- Eliminate GPU underutilization
- Scale across exchanges reliably
- Reduce manual IT burden
- Eliminate idle GPU and CPU waste

Storage Built for Always-On, High-Stakes Market Platforms

To accelerate financial growth, **81% of banking CFOs are projecting an increased spend on AI uses in 2026.** For high-frequency trading firms, faster AI models are now required to outtrade the competition during peak market windows because a single latency spike eliminates profit potential. Infrastructure built around ultra-low latency and maximum performance results in microsecond-scale execution and probabilistic throughput that is the difference between trade captures and misses. High frequency traders know that faster infrastructure is now a competitive edge and modernizing the storage stack helps **make more money**.

Co-located at exchanges, HFT firms need storage that holds microsecond-scale latency under concurrent live order books, market data feeds, and risk model scoring. **NeuralMesh Augmented Memory Grid™** sustains that performance during the peak windows when trading volume is highest and profit is made or lost. On-prem deployment keeps data adjacent to the trading engines that consume it, so firms move from one exchange to the next on the same proven blueprint.

> *"Our existing Network Attached Storage was too slow to effectively cover analytics workloads. **WEKA proved to be 10× faster** and allowed for seamless backup to our object store."*
>
> — WEKA Capital Market Customer