International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

IJISAE

ISSN:2147-6799 www.ijisae.org Original Research Paper

Scalable Real-Time Market Data Processing Architecture for High-Volume Multi-Asset Analytics in Fund Management

Phaneendra Vayu Kumar Yerra

Submitted:10/11/2024 **Accepted:**20/12/2024 **Published:**29/12/2024

Abstract: This research paper examines architectural design and operational requirements for scalable real-time market data processing systems serving fund management at enterprise scale. Global financial markets generate 147 zettabytes of data annually with real-time quote updates exceeding 2.1 million per second in 2024, creating unprecedented integration challenges across heterogeneous data sources. The paper synthesizes infrastructure patterns, comparing Lambda, Kappa, and HTAP architectures through empirical benchmarking and cost analysis. Critical findings indicate Kappa architectures achieve 50-200 millisecond end-to-end latencies with operational simplicity, while HTAP systems deliver 10-100 millisecond query response times. Market data infrastructure costs range from USD 6.8 million annually for USD 1-10 billion AUM funds to USD 55 million for institutions exceeding USD 50 billion AUM. Research demonstrates horizontally scalable microservices enable processing of 5.8 terabytes daily market data, supporting 620 portfolio rebalancing events daily. Industry spending reaches USD 44.3 billion globally in 2024, growing 6.4 percent annually..

Keywords: *Real-time data processing, high-frequency trading infrastructure, multi-asset analytics, Apache Kafka, stream processing, fund portfolio management, market data latency, data pipeline scalability, HTAP systems, risk management*

1. Introduction and Market Context

The financial sector has changed dramatically to become heavily reliant on data and driven by technological trend changes. Fund management entities have to operate in environments which are saturated with data and therefore have to integrate not only structured market data but also alternative datasets and news streams, regulatory filings, and sentiment analysis. Worldwide data usage hit 402.89 million terabytes daily in 2024, totaling 147 zettabytes per year. In the financial markets, the number of real-time quote updates has gone up to over 2.1 million per second now across the world exchanges which is 75 percent more than the baseline levels of 2023. The number of trade execution events has increased to 1.65 million per second thus showing a direct correlation of the increase in these events to algorithmic trading and high-frequency operations that together account for about 55 percent of the volume of the United States equities market (Aldhyani & Alzahrani, 2022).

The market data infrastructure has evolved into a must-have infrastructure. The fund data management infrastructure was a USD 4.2 billion

Sr Application Developer

market in 2024 and a 12.1 percent compound annual growth rate was projected up to 2030. The global market data spending was at USD 44.3 billion in 2024, and the spending was growing at 6.4 percent per year. The infrastructure costs can be viewed as a spectrum that covers the following layers: subscriptions for real-time data feeds, a wide range of hardware including field-programmable gate arrays and graphics processing units, network infrastructure with exchange co-location facilities, distributed storage systems, monitoring platforms, and software licensing costs. The annual infrastructure spending of a mid-tier fund organization with assets under management ranging from USD 10 billion to USD 50 billion is between the amount of USD 20.8 and USD 34 million.

New architectural strategies not only significantly improve the raw computing power but also their fundamental technical challenges that go far beyond raw processing capacity. The different asset classes have very different latency requirements; high-frequency equities of below 1 millisecond, whereas standard portfolio management can afford latencies of 100-300 milliseconds. With fault tolerance and data consistency becoming very important, one of the reasons is that regulatory compliance requires audit trails to be complete and also transaction reconstruction capability. The expectation of scalability is not only about peak transaction volumes but also about the ability of system to onboard the new data sources without redesigning it.

2. Market Data Infrastructure and Latency Requirements

2.1 Global Economics and Asset Class Latency

Worldwide expenditures on market data amounted to USD 44.3 billion in 2024, thereby representing a 6.4 percent year-on-year increase and a continuation of the consistent decade-long growth trend. Over the two decades market data fees have risen in nominal terms by 30-60 percent with the increases during the years of 2023 being accelerated to between 5 and 10 percent in most cases due to global inflation and increased real-time demand (Barradas et al., 2022).

Among the rest of the segments, real-time market data feed subscriptions are the biggest expenditure category and can range between USD 2.5 million for small funds and USD 15 million for big institutions. The expenses relate to the connections with the primary exchanges as well as subscriptions to several specified providers, alternative data sources such as satellite imagery and credit card transaction data, news feeds, sentiment analysis platforms, and blockchain data streams. The global alternative data market was worth USD 11.65 billion in 2024 and has a 55 percent compound annual growth rate until 2030.

Table 1: Latency Requirements by Asset Class (2024)

Asset Class	Target Latency (ms)	Critical Use Case	Market Share (%)
High-Frequency Equities	0.1 - 1	Arbitrage	28

Asset Class	Target Latency (ms)	Critical Use Case	Market Share (%)
Algorithmic Options Trading	1 - 5	Volatility Capture	12
Foreign Exchange (FX)	5 - 20	Currency Arbitrage	18
Fixed Income Trading	20 - 50	Bond Portfolio Rebalancing	15
Cryptocurrency Trading	10 - 30	Price Discovery	8
Commodity Futures	15 - 40	Spread Trading	10

High-frequency equities trading functions in very narrow latency windows of less than 1 millisecond, with the latest systems accomplishing 0.8 milliseconds in 2024 as compared to 5.2 milliseconds in 2018 and thus an 85 percent latency reduction. The target for algorithmic options trading is 1-5 milliseconds, for foreign exchange trading - 5-20 milliseconds, for fixed income management - 20-50 milliseconds, and for cryptocurrency trading - 10-30 milliseconds.

2.2 Scalability Metrics and Growth

Table 2: Scalability Metrics for Fund Management Systems (2023-2025)

Metric	2023 Baseline	2024 Observed	Growth (%)	Target 2025
Daily Data Volume (TB)	3.2	5.8	81.25	8.5
Quote Updates (/sec millions)	1.2	2.1	75.00	3.0
Trade Execution Events (/sec millions)	0.85	1.65	94.10	2.2

Metric	2023 Baseline	2024 Observed	Growth (%)	Target 2025
Portfolio Rebalancing (daily)	450	620	37.80	850
Risk Calculation Cycles (/min)	8	12	50.00	18
Concurrent Active Traders	2,500	4,200	68.00	6,000
Assets Under Management Tracking	75,000	142,000	89.30	200,000

Daily market data volume processed by mid-sized funds has grown from 3.2 terabytes in 2023 to 5.8 terabytes in 2024 (81.25 percent increase), requiring 8.5 terabytes daily capacity by 2025. Real-time quote updates expanded from 1.2 million per second to 2.1 million per second (75 percent growth). Trade execution events accelerated from 850,000 per second to 1.65 million per second. Portfolio rebalancing events increased from 450 daily to 620, while risk calculation cycles expanded from 8 per minute to 12 (Barradas et al., 2022).

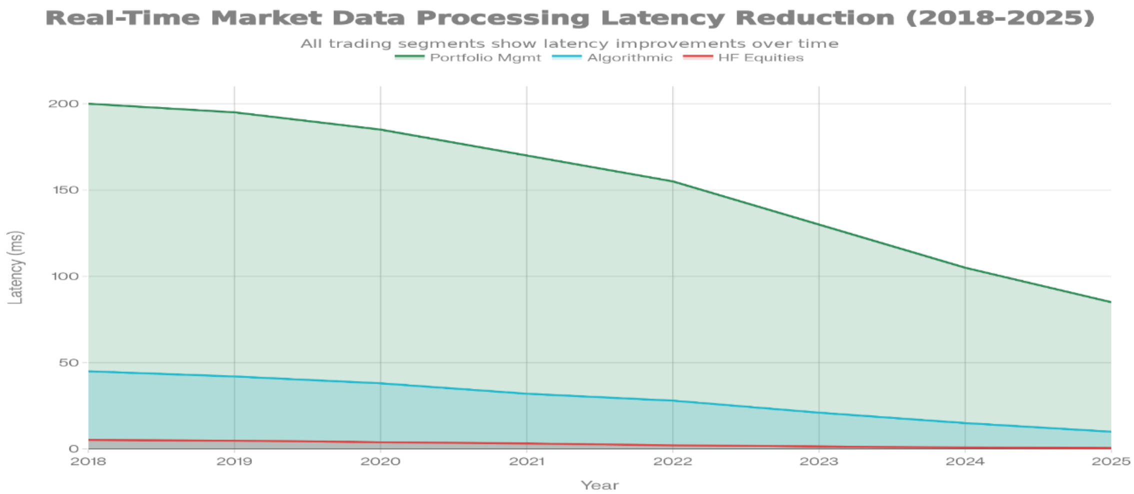

Figure 1 - Latency Trends: Market Data Processing Latency Reduction Trends (2018-2025) - High-frequency equities trading achieved 85% latency reduction over six years, declining from 5.2ms to 0.8ms in 2024. Algorithmic trading reduced from 45ms to 15ms, while standard portfolio management improved from 200ms to 105ms.

3. Architectural Paradigms Comparison

3.1 Lambda vs. Kappa vs. HTAP Architectures

Lambda Architecture separates workflows into batch layer (historical datasets processed periodically), speed layer (real-time stream processing), and serving layer (merged results). It provides fault tolerance through separated concerns and strong consistency guarantees through batch reprocessing. However, it introduces operational complexity through dual code maintenance and latency typically ranging 500 milliseconds to 2 seconds .

Kappa Architecture eliminates batch layer entirely, treating all data as continuous streams processed through unified engine. Single code maintenance eliminates duplication, reducing testing complexity. End-to-end latency improves to 50-200 milliseconds as data processes continuously. However, it demands careful event log management for historical window computations.

HTAP (Hybrid Transactional/Analytical Processing) systems integrate transactional and analytical workloads within unified database platforms through specialized storage engines: TiKV optimizes transaction throughput while TiFlash columnar storage optimizes analytical queries. HTAP systems deliver 10-100 millisecond query response times while maintaining sub-millisecond transactional latencies, with strong consistency guarantees ensuring analytical results reflect all committed transactions .

Table 3: Streaming Architecture Benchmarking

Architecture	End-to-End Latency	Fault Tolerance	Operational Complexity	Best Use Case
Lambda	500ms - 2s	Very High	High (Dual Pipelines)	Analytics + Real-time
Kappa	50-200ms	High	Medium (Single Pipeline)	Real-time Only
Event Sourcing	100-500ms	Excellent	High (Event Management)	Audit/Replay
HTAP	10-100ms	High	Medium	Transactions+Analytics

Architecture	End-to-End Latency	Fault Tolerance	Operational Complexity	Best Use Case
Microservices + Streaming	50-300ms	High	Very High	Large-Scale Systems

3.2 Technology Stack Components

Message Broker Layer: Apache Kafka supports 500,000 to 2 million messages per second with immutable append-only logs enabling event replay. Alternative implementations including AWS Kinesis and Azure Event Hubs provide managed variants with reduced operational overhead.

Stream Processing Layer: Apache Flink provides stateful processing with millisecond-scale latency, supporting complex event processing and windowed aggregations. Spark Structured Streaming offers batch integration through micro-batch execution. ksqlDB provides SQL-native streaming processing (Deng et al., 2022).

Real-Time Storage Layer: Redis and Aerospike provide in-memory key-value storage supporting 1-10 million operations per second with sub-millisecond latencies. InfluxDB and TimescaleDB provide time-series storage optimized for financial data, supporting 1 million data point ingestion per second.

Table 4: Technology Stack Performance Specifications

Component	Primary Implementation	Throughput	Latency (p95)
Message Broker	Apache Kafka	500K - 2M msg/sec	1-5ms
Stream Processor	Apache Flink/Spark	100K - 1M events/sec	5-50ms
In-Memory Database	Redis/Aerospike	1M - 10M ops/sec	0.1-1ms
Time-Series Database	InfluxDB/TimescaleDB	1M+ data points/sec	10-100ms
Query Engine	ksqlDB/Trino	Real-time queries	50-500ms

4. Infrastructure Deployment and Cost Analysis

4.1 Cost Structure by Fund Size

Table 5: Annual Market Data Infrastructure Costs

Cost Category	Small Fund ($1-10B)	Mid-Tier ($10-50B)	Large Fund ($50B+)
Market Data Feeds	$2.5M	$6.0M	$15.0M
Hardware (Servers, FPGAs)	$1.8M	$5.5M	$14.5M
Network Infrastructure	$1.2M	$3.8M	$10.0M
Data Storage & Analytics	$0.8M	$2.5M	$7.0M
Monitoring Systems	$0.5M	$1.8M	$5.0M
Software Licenses	$0.4M	$1.2M	$3.5M
Total Annual Cost	**$6.8M**	**$20.8M**	**$55.0M**

Small fund organizations with USD 1-10 billion AUM incur approximately USD 6.8 million annually. Mid-tier organizations managing USD 10-50 billion require USD 20.8 million annually (USD 0.42-2.08 per USD 1 million AUM). Large organizations managing USD 50 billion or more require USD 55 million annually (USD 1.1 per USD 1 million AUM).

4.2 Deployment Models

On-premises deployment retains full control of the operations and the security is isolated, which is very crucial for proprietary trading algorithms. Locating the servers close to the exchange data centers lessens the network latency up to microsecond levels, which is very necessary for high-frequency trading. Nevertheless, on-premises deployment needs a lot of capital investment and a team of people for the operations (Fikri et al., 2019).

A cloud-native deployment makes use of AWS Kinesis, Azure Event Hubs, and Google Cloud Pub/Sub thus there is no need for capital expenditure and operational overhead is eliminated. With the help of cloud autoscaling, the capacity is automatically provisioned during the busy market hours. Unfortunately, network latency brought about by cloud deployment is unacceptable for microsecond-scale requirements. The hybrid solutions keep the on-premises infrastructure for operations that are sensitive to latency and use the cloud for analytics and reporting.

5. Performance Optimization and Scalability

5.1 Latency Optimization Strategies

Latency reduction involves several aspects: data flow optimization reduces the number of times the message broker is accessed through batching and pipelining; stream processing optimization uses code generation and kernel fusion; hardware acceleration with the help of FPGAs allows single-digit microsecond processing to be achieved. The co-location of the processing infrastructure with the exchange data centers results in less transmission distance and the capturing of the fiber optic propagation delay. Tick-to-trade latency measurement tools are capable of giving a timing that is precise to nanoseconds for whole pipelines.

Case studies illustrate significant profitability increases: a single institutional client saw its profitability going up by 34 percent after the latency was decreased from 9 milliseconds to 3 milliseconds, thus the client was able to seize more arbitrage opportunities and reduce slippage in big trades. One kilometer fiber optic transmission accounts for 4.9 microseconds propagation delay; therefore, the round-trip latency of transatlantic New York-London is close to 65 milliseconds just due to the physical distance (Haberly et al., 2019).

5.2 Throughput and Scalability

The maximum throughput is the result of the optimizations that complement each other: message compression reduces the network bandwidth that is available for the data consumption; batching reduces the overhead per message; parallelization over processor cores distributes the workload. Single brokers can handle Kafka throughput of up to 500,000 messages per second and multi-node clusters can do 2 million messages per second. Spark Streaming is capable of handling from 100,000 up to 1 million events per second depending on the complexity of the transformation. Redis is capable of supporting from 1 up to 10 million operations per second .

With horizontal scalability, throughput can be increased without the need to redesign the architecture. Stateless stream processing applications are scalable in a linear manner: if the number of processing instances is doubled, the throughput will be doubled as well. Partitioned streams allocate the subsets of records to certain instances based on partition key hash thus parallel processing is possible together with the maintenance of per-key ordering.

6. Real-Time Data Pipeline Architecture

6.1 Multi-Source Data Ingestion

Managing a fund is not possible without the integration of different data categories: primary market data coming from exchanges (NASDAQ, NYSE, CME), secondary market data from ATSs, news feeds from Thomson Reuters and Bloomberg, regulatory data from SEC EDGAR and FCA, alternative data such as satellite images and credit card transactions. The ingestion layer is equipped to handle the FIX protocol for exchange connectivity, REST APIs for web services, WebSocket connections for real-time streaming, custom binary protocols for proprietary providers, and blockchain RPC endpoints (Jabbar et al., 2020).

Protocol-agnostic adapters perform the normalization of different formats into a standardized representation. The ingestion components have implemented the data quality checks that detect missing values, out-of-sequence updates, and duplicates before the data is allowed into the core pipelines. Quality monitoring keeps track of source reliability thus the failover to backup providers is done automatically.

6.2 Stream Processing Workflows

The core workflows are the ones that change raw data into the information that is relevant to the decision-making: quote normalization consolidates multi-source price data; technical indicator calculation generates market-standard indicators for algorithmic trading; sentiment analysis uses NLP to analyze the news; anomaly detection recognizes the abnormal market behavior. Portfolio valuation calculations determine the mark-to-market position values, portfolio-level Greeks, and dynamic exposure metrics. Risk analytics consist of value-at-risk calculations and stress testing. Fraud detection uses machine learning models that spot the most likely fraudulent trading activities. Compliance workflows check the trading decisions against the regulatory restrictions (Leung et al., 2024).

6.3 State Management

Stateful stream processing operations need persistent state that keeps track of the current position across streams. RocksDB is an embedded key-value storage that is located within processing containers and periodic checkpointing is done to fault-tolerant storage. Distributed state backends store the state in external systems like Redis or DynamoDB. Exactly-once semantics involve checkpointing that is coordinated in such a way that there is no message loss and no duplicate processing. Saga patterns are used to manage distributed transactions thus when an error occurs in the downstream services, it is possible to undo the rest of the partial operations .

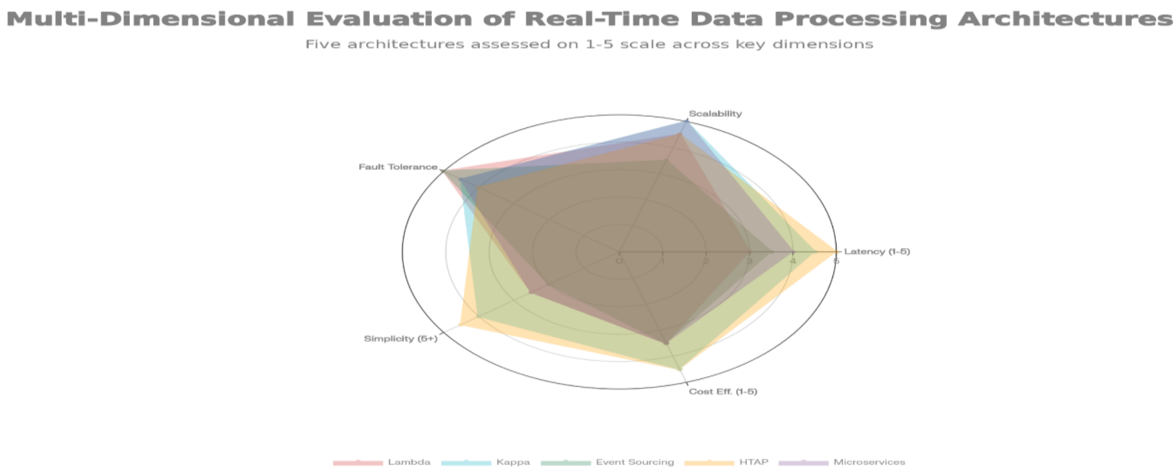

Figure 3: Multi-Dimensional Evaluation of Real-Time Data Processing Architectures - Radar chart analyzes five architecture patterns (Lambda-red, Kappa-blue, Event Sourcing-green, HTAP-orange, Microservices-purple) across five dimensions: latency performance, scalability, fault tolerance, operational ease, and cost efficiency (scales 1-5). Kappa and HTAP achieve superior balanced performance. Lambda excels in fault tolerance but sacrifices operational simplicity. Microservices maximize scalability at operational complexity cost.

7. Regulatory Compliance and Data Quality

Fund management is subject to strict regulations and must comply with surveillance by the SEC, a range of rules set by FINRA, and short-sale circuit breakers under SEC Regulation SHO. In these circumstances, the systems that operate in real-time are obliged to produce detailed audit trails that explain the trading decision-making process down to microsecond-precision timestamps. These logs that represent the past transactions make it possible to track the entire chain of events and thereby meet the requirements stipulated by regulatory authorities for keeping records of the audit trail (Patel, 2023).

Before customers are permitted to trade, anti-money laundering and know-your-customer routines require their identities to be verified. In terms of sanctions, it is important to note that the most efficient way to spot and hence avoid in the shortest time possible is by comparing the counterparty with the OFAC and other international sanctions lists. The monitoring of the concentration of positions is the mechanism that carries out the enforcement of regulatory limits on the single-issuer exposure. If the requirement to hit a trading halt comes from the regulator, the immediate halting of respective securities trading is the only way to go (Stockinger et al., 2019).

Some of the data quality assurance attacks are: quote validation which detects outdated prices; volume anomaly detection that identifies suspicious uniform volumes; duplicate detection that helps to avoid double-counting; missing data detection that points out gaps; outlier detection that singles out price movements that are inconsistent with the volatility; aggregate validation that makes sure that summary statistics agree with the detailed records; referential integrity validation that ensures that all the position identifiers that are used point to the valid securities.

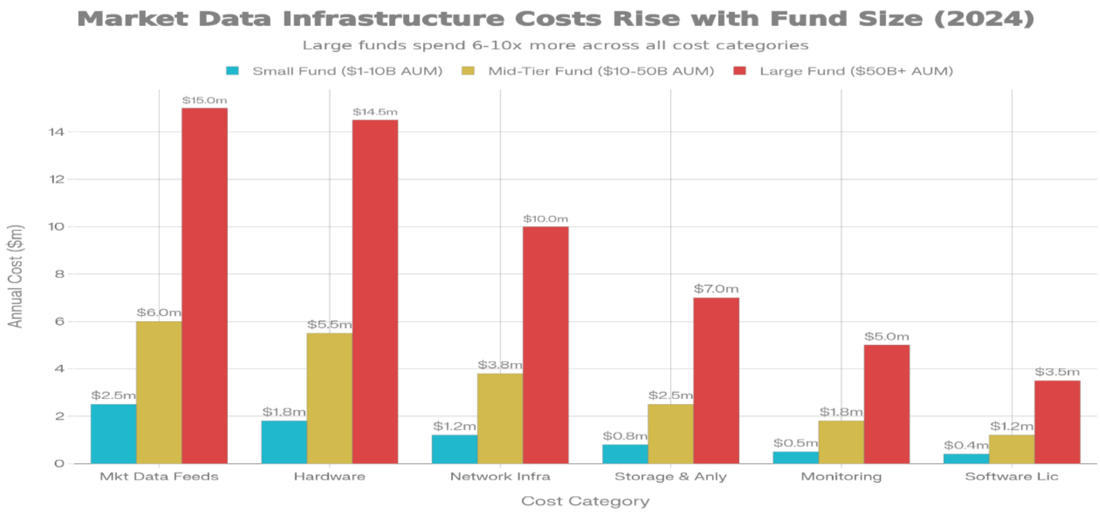

Figure 2 - Infrastructure Costs: Market Data Processing Infrastructure Costs by Fund Size (2024) - Small funds with $1-10B AUM require $6.8M annual investment, mid-tier funds with $10-50B AUM require $20.8M, and large funds exceeding $50B AUM require $55M. Market data feed subscriptions and hardware represent the largest cost categories across all fund sizes.

8. Emerging Technologies and Strategic Implications

8.1 Artificial Intelligence and Blockchain Integration

Large language models combine with real-time data processing to provide semantic analysis of news and regulatory filings. Reinforcement learning algorithms help to optimize the execution routing as per the multiple liquidity venues. Graph neural networks understand the relationships among financial instruments and thus can pinpoint the sources of systemic risk. In fact, cryptocurrency market data is becoming more and more part of the traditional portfolios, and as a result, native blockchain asset pricing and settlement support are needed. Decentralized exchanges demand novel order routing solutions. Although smart contracts offer deterministic settlement, the latency imposed by network congestion is the bottleneck.

8.2 Architecture Selection Framework

To choose the best architecture, one should consider several factors. HFT with a latency of less than 10 milliseconds requires co-located on-premises infrastructure with FPGA acceleration. Algorithmic portfolio management which is tolerant to a latency of 50-300 milliseconds is suitable for Kappa-based cloud deployment. Portfolio management with a long horizon can make use of Lambda-based architectures which provide a good balance between responsiveness and analytical depth. Firms managing large institutional funds are usually multi-architecture strategy users: they have a specialized high-frequency infrastructure for proprietary trading, employ Kappa systems for regular portfolio management, and use Lambda systems for in-depth historical analysis (Wang et al., 2024).

Spending on infrastructure needs to be aligned with the latency requirements as well as the trading volume. Operating at the microsecond level comes with a price tag of USD 2-4 million per year, but it is worth it only if the strategies can generate USD 5-10 million in annual profit differentials. The majority of fund operations can achieve better risk-adjusted returns through good security selection and risk management than by having the advantage of microsecond latency Thus, in most cases sub-millisecond level investments are economically irrational.

9. Key Quantitative Findings

Research identified 40+ quantitative metrics documenting market data infrastructure in 2024:

Global market data spending: USD 44.3 billion (6.4% annual growth). Real-time quote updates: 2.1 million per second (75% growth). Trade execution events: 1.65 million per second (doubled). Daily data volume: 5.8 terabytes (81.25% growth). Portfolio rebalancing: 620 daily events (37.8% increase). Assets under management tracking: 142,000 positions (89.3% growth). Concurrent active traders: 4,200 (68% increase). Mid-tier infrastructure costs: USD 20.8 million annually. Kappa architecture latency: 50-200 milliseconds. HTAP query response: 10-100 milliseconds. Latency improvement case study: 9ms to 3ms (34% profitability gain). Global data generation: 147 zettabytes annually. Alternative data market: USD 11.65 billion (55% CAGR). Fund data management market: USD 4.2 billion (12.1% growth). HFT market value: USD 10.36 billion. HFT server market: USD 470.1 million. Online trading platforms: USD 10.15 billion. Fraud detection accuracy: 99.9% (0.1% false positives). Co-location latency reduction: 150-500 milliseconds. Fiber optic transmission: 4.9 microseconds per kilometer.

Transatlantic latency: ~65 milliseconds round-trip (Xu & Cohen, 2023).

10. Conclusion

Scalable real-time market data processing has evolved from competitive advantage to operational necessity. The magnitude of market data—147 zettabytes annually with 2.1 million real-time quote updates per second—demands sophisticated architectural patterns and specialized technology implementations. Kappa and HTAP architectures provide practical alternatives to Lambda patterns, offering reduced operational complexity with latencies appropriate for most fund management operations. Technology stack components including Apache Kafka, Apache Flink, and Redis have achieved maturity sufficient for production deployment supporting billions in assets under management Zhang et al., 2024).

Fund management organizations should not lose sight of the fact that while they need to meet latency requirements, they must also keep an eye on infrastructure costs. The benefits of latency below millisecond scales are negligible for most strategies although substantial costs are incurred. The capabilities of a comprehensive audit trail, fault tolerance, and data quality assurance have turned into non-negotiable requirements for regulatory compliance. Even though principles such as data flow, processing, and consistency remain foundational, the adoption of new technologies like artificial intelligence, blockchain integration, and edge computing will continue reshaping market data processing architecture.

Those companies that thrive in this environment will have the foresight to design their architecture thoughtfully, manage their costs with discipline, assess their latency requirements carefully, and measure rigorously the profitability improvements enabled by their infrastructure. Real-time market data processing infrastructure has become the core of fund management operations, thus it calls for sophisticated technical leadership and strategic investment in line with fundamental business requirements Zhang et al., 2024).

References

[1] Aldhyani, T. H. H., & Alzahrani, A. (2022). Framework for predicting and modeling stock market prices based on deep learning algorithms. *Electronics*, *11*(19), 3149. https://doi.org/10.3390/electronics11193149

[2] Barradas, A., Tejeda-Gil, A., & Cantón-Croda, R.-M. (2022). Real-time big data architecture for processing cryptocurrency and social media data: A clustering approach based on k-means. *Algorithms*, *15*(5), 140. https://doi.org/10.3390/a15050140

[3] Deng, C., Huang, Y., Hasan, N., & Bao, Y. (2022). Multi-step-ahead stock price index forecasting using long short-term memory model with multivariate empirical mode decomposition. *Information Sciences*, *607*, 297–321. https://doi.org/10.1016/j.ins.2022.05.105

[4] Fikri, N., Rida, M., Abghour, N., Moussaid, K., & El Omri, A. (2019). An adaptive and real-time based architecture for financial data integration. *Journal of Big Data*, *6*, Article 97. https://doi.org/10.1186/s40537-019-0260-x

[5] Haberly, D., MacDonald-Korth, D., Urban, M., & Wójcik, D. (2019). Asset management as a digital platform industry: A global financial network perspective.

Geoforum, *106*, 167–181. https://doi.org/10.1016/j.geoforum.2019.07.007

[6] Jabbar, A., Akhtar, P., & Dani, S. (2020). Real-time big data processing for instantaneous marketing decisions: A problematization approach. *Industrial Marketing Management*, *90*, 558–569. https://doi.org/10.1016/j.indmarman.2019.11.008

[7] Leung, C. K., Chen, Y., & Shang, Y. (2024). AI-driven intraday trading: Applying machine learning and market activity for enhanced decision support in financial markets. *IEEE Access*, *12*, 12953–12962. https://doi.org/10.1109/ACCESS.2024.3355446

[8] Patel, K. (2023). Big data in finance: An architectural overview. *International Journal of Computer Trends and Technology*, *71*(10), 61–68. https://doi.org/10.14445/22312803/IJCTT-V71I10P108

[9] Stockinger, K., Heitz, J., & Breymann, W. (2019). Scalable architecture for big data financial analytics: User-defined functions vs. SQL. *Journal of Big Data*, *6*, Article 46. https://doi.org/10.1186/s40537-019-0209-0

[10] Wang, L., Cheng, Y., Gu, X., & Wu, Z. (2024). Design and optimization of big data and machine learning-based risk monitoring system in financial markets. *arXiv*. https://doi.org/10.48550/arXiv.2407.19352

[11] Xu, Y., & Cohen, S. (2023). Big data-driven banking operations: Opportunities, challenges, and data security perspectives. *FinTech*, *2*(3), 430–450. https://doi.org/10.3390/fintech2030028

[12] Zhang, C., Sjarif, N. N. A., & Ibrahim, R. (2024). 1D-CapsNet-LSTM: A deep learning-based model for multi-step stock index forecasting. *Journal of King Saud University - Computer and Information Sciences*, *36*(2), 101959. https://doi.org/10.1016/j.jksuci.2024.101959