Beeks

Introduction to Edge Intelligence

April 24, 2025

Build.Connect.Analyse.

Table of contents

Customer Document Name	Introduction to Edge Intelligence
Customer Document ID	BKDA014W

Document control

SCOPE

Technological advances in AI and high-speed processing of very high volume data via GPU technology have presented new opportunities to extract data and generate AI-powered analytics close to the data source, and achieve the holy grail of trading insights - exceptionally low-latency and context-aware analytics.

This paper focuses specifically on trading environments and examines the role of edge analytics, the ways in which AI transforms edge analytics, and the many advantages of using AI-powered analytics in edge computing to extract insights that will transform network performance.

AUDIENCE

The intended audience for this document is any business or technical stakeholder with an interest in learning about AI-powered edge analytics and Beeks Market Edge Intelligence solution.

STYLING CONVENTIONS USED IN THIS DOCUMENT

When this document refers to *something out of the ordinary*, it adopts a different formatting (of which this is an example). Users who *only want ordinary things* can ignore these references.

Proximity Cloud > Rack Overview. This indicates a sequence of expandable menu items to follow to reach the desired item (Rack Overview in this case).
Each > indicates that the item is a sub-item of the one preceding it.

FEEDBACK

We welcome feedback on our documentation. Please report suggestions, omissions and errors to documentation@beeksgroup.com .

RELATED DOCUMENTS

Document name	Beeks document reference	Description
Analytics Concepts Guide	BKDA001	Our introduction to analytics in Beeks, covering all concepts and architecture.
Analytics Core Data Feed Guide	BKDA003	This document describes the Core Data Feed in Beeks Analytics, which enables you to stream metrics for visualisation and analytics. Its purpose is to explain the supported data streams, their use, format, and configuration.
Beeks Analytics vs Corvil: A Strategic Comparison	BKDA015	This document compares Beeks Analytics functionality and performance to those of Corvil.

Documents that do not have hyperlinks can be obtained by contacting your Beeks account representative.

VERSION HISTORY

Version	Description	Saved on
V2	Second version. Split into separate documents: Introduction to Edge Intelligence (BKDA014W), and White Paper Edge Intelligence (BKDA014N).	Apr 24
v1	First issued version.	Apr 14

Executive Summary

As trading environments become increasingly complex and data volumes explode, the financial sector faces a pivotal challenge: how to derive actionable insights at the lowest possible latency without overburdening central systems. Advances in GPU computing and AI make it possible to analyze data directly at the network edge—where it originates—bringing capital markets closer than ever to real-time intelligence.

This paper focuses on the role of edge analytics in trading environments, demonstrating how AI transforms edge intelligence and the many advantages of deploying AI-powered analytics within an edge computing framework to optimise network performance.

The paper also describes Beeks Market Edge Intelligence, which combines statistical baselines with contextual machine learning models to provide a unique and nuanced view of market data health. By detecting anomalies early - before they become visible at the application layer - trading firms can respond faster, reduce operational risk, and maintain the quality of their trading infrastructure in the most demanding market conditions.

Beeks Market Edge Intelligence delivers real-time insights by integrating AI directly into the trading infrastructure. This transforms how trading firms monitor, diagnose, and optimise performance - before issues arise.

Unlike closed, appliance-based platforms, Beeks embraces an open, modular architecture. The platform delivers real-time anomaly detection, forecasting, and predictive maintenance via GPU-accelerated models, while giving firms full ownership and accessibility over their data. Integration with open standards like Kafka and QuestDB ensures seamless ingestion, portability, and high-speed access for machine learning pipelines.

Market Edge Intelligence offers:

- **Low latency** AI-driven analytics deployed at the edge.
- **Direct access to GPU-capable servers** for high-speed model execution.
- **Seamless API integration** for GPU-accelerated models.
- **Proactive anomaly detection and forecasting** to minimise disruptions.
- **A modular, open architecture** to ensure full data accessibility and interoperability.

For trading firms, adopting AI-driven edge analytics is a game-changer—reducing latency, enhancing operational visibility, and ensuring an unmatched trading advantage.

Introduction

The need for speed

In today's high-frequency trading landscape, stakeholders expect rock-solid connectivity and exceptional network performance. The ability to capture, process, and analyse vast volumes of data at speed to detect performance loss and resolve issues fast is a key differentiator for trading venue providers and their clients.

However, firms face many challenges that are not adequately addressed by today's available technology:

- Quants review financial data but this often misses information about the performance over the network, and it misses latency information.
- Certain quant analysis is not fast enough to feed back into live trading decisions, or is not accessible by operational teams who need to use it to inform real-time intraday decisions about connectivity, infrastructure, counterparties, etc.

Recent advances in AI and GPU technology have yielded many innovations that attempt to meet this challenge. Of these, AI-powered edge analytics offers the most promise in addressing these issues.

This paper describes how the Beeks Analytics architecture contributes to and enhances customers' existing AI strategies; and then expands on how the Beeks Market Edge Intelligence product will supplement this further by:

- delivering low latency AI-powered analytics closer to the source data.
- providing direct access to GPU-capable servers in colo sites.
- providing API access to GPU-accelerated hosted models.

Market Edge Intelligence will be available as part of Beeks Analytics. It will also be available as a standalone product to operate on customer's own data that they collect from the edge but struggle to analyse.

How Beeks Analytics already supports client AI initiatives

Beeks Analytics is a suite of products that records and analyses latencies at network and application level, delivering exceptional, real-time insight into performance so that our clients can build, maintain, and deliver a trading environment that gives their customers the best trading experience.

Beeks Analytics provides unrivalled visibility that enables you to track and analyse the real-time performance of every single message traversing business critical processes. Identify outliers and bottlenecks, understand capacity issues, verify new system roll-out performance, and ultimately ensure delivery of a consistent and high-quality trading environment.

SOME EXAMPLE USE CASES THAT ARE SOLVED BY BEEKS ANALYTICS:

- Do I have performance problems with any of my trading counterparties?
- Is the problem inside my infrastructure or outside my infrastructure?
- I need an overview of the flows going across my network to assist with troubleshooting and capacity.
- I need to maintain Service Level Agreements within my trading system for my customers.
- I need to exactly reproduce a period of sub-par performance in my trading system to try some possible solutions.

The goals behind the Beeks Analytics architecture are as follows:

- To operate with an ***open architecture***, which provides the benefit of multiple high volume integration points with your organisation's own systems.
  - Allowing your organisation to fully own the data produced by the Analytics system.
  - Enables the organisation to run the Analytics software on their own hardware.
- To allow ***open scaling***, which allows the capacity of the system to handle load to scale up with commercially available server hardware improvements.
- To be modular and licenced to support ***open consumption***.
  - This means that, for example, if you only need the high performing VMX-Capture layer and don't require the in-depth analytics that

the VMX-Analysis layer provides, we'll ensure the software is licensed and priced accordingly.
- It also means that we're transparent about the drivers of our pricing - core count required for the analysis, which we make clear in our transparent performance metrics.

- To support ***open data*** by ensuring that you have full access to, and control over, all monitoring data generated by the platform.
  - Rather than locking data into proprietary interfaces, our architecture emphasises direct and flexible data accessibility.
  - Our Advanced Configurable Decoder™ (ACD) ensures an agile, lower-cost way to monitor internal messaging data on the wire.
  - The Kafka-based Core Data Feed provides robust scalability and fault tolerance even under high message rates to provide lossless, low latency data streams. Use of Kafka frees you from vendor lock-in, since Kafka connectors exist for virtually every modern data processing framework, making it straightforward to combine Beeks Analytics output with your broader enterprise data.

THE ROLE OF THE CORE DATA FEED

The Beeks Analytics Core Data Feed (CDF) provides a powerful foundation for AI development by delivering a consistent, high-volume stream of normalized trading data in real time. Because the CDF unifies multiple data streams into one central analytics pipeline, it removes the complexities that typically arise when data is scattered across numerous capture points. This significantly reduces integration overhead, making it easier to feed large, high-quality datasets into AI platforms. In short, the CDF's consolidated approach unlocks new possibilities for machine learning projects, because machine learning exports do not have to juggle multiple streams or worry about synchronisation issues when training algorithms.

Moreover, CDF's flexible configuration ensures that AI models receive only the necessary data, preserving compute resources and accelerating the training loop. Since developers can fine-tune which metrics or decoded messages get passed into Market Edge Intelligence or to the client's own AI models, they can prevent the "noise" of unnecessary data. This efficiency not only bolsters model performance but also reduces the risk of processing bottlenecks or system failures. As a result, financial institutions can build and refine high-performance trading analytics faster, with fewer hurdles to real-world deployment.

For more about the Core Data Feed, see the *Analytics Core Data Feed Guide* or the summary in the appendix of this document (see page 32).

CASE STUDY: DISTRIBUTED CAPTURE, TELEMETRY AND DECODE IN THE NEW JERSEY TRIANGLE

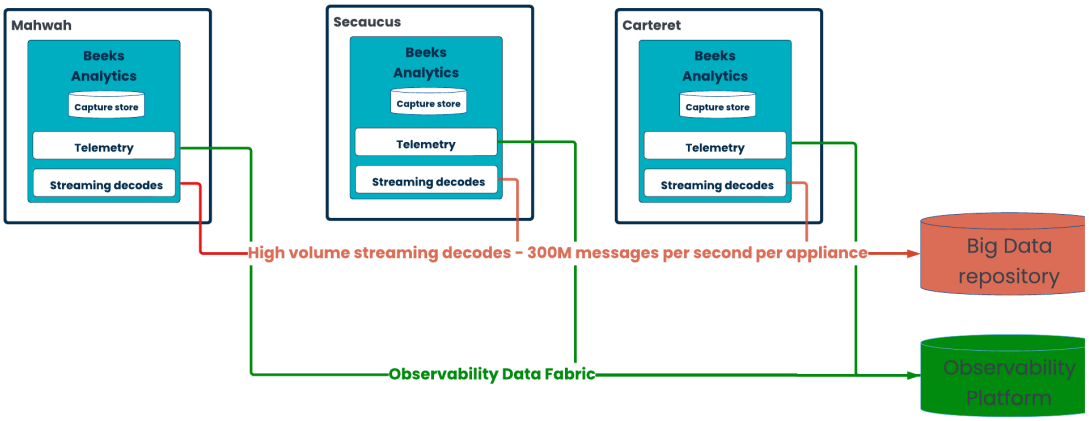

Figure 1 Observability data fabric.

The New Jersey Triangle cluster of datacentres in Secaucus, Weehawken, and Carteret are an interesting example of how the architecture of Beeks Analytics supports firms central AI initiatives. These datacentres house multiple exchange matching engines and provide critical infrastructure for high-frequency trading (HFT) firms.

The U.S. stock markets produce a vast amount of market data primarily due to their size, liquidity, and competitive structure. Regulations such as Regulation NMS (National Market System) encourage competition among multiple exchanges, each generating and disseminating its own price quotations and trades. Furthermore, the depth and maturity of the derivatives markets, combined with the global significance of the U.S. economy, attract high participation volumes and rapid quote updates—ultimately contributing to an enormous amount of market data traffic.

As of 2024, NYSE/ICE and Nasdaq collectively control over 90% of U.S. equity trading volume, with remaining activity spread across 14 other SEC-regulated exchanges like CBOE and IEX.

Part of the reason that there's so much trading on these exchanges is that, since the Unlisted Trading Privileges (UTP) Act in 1994, stocks have been allowed to trade on multiple exchanges. That meant the primary listing exchange was no longer the only exchange on which a stock could trade.

What benefit is there to the investor of having a choice of exchanges to trade on? Different exchanges might offer different fee structures, or slightly different algorithms for matching buyers and sellers of a particular stock. One of these might suit your trading strategy more than others. The US market structure tries to balance the needs of the institutional investor to optimise execution costs through 'exchange selection' strategies and fee competition between exchanges, whilst benefiting retail traders by

mandating order routing to the particular venue which is offering the best prices at that time.

Although the US equity trading market is diverse in terms of the number of venues on which stocks can be traded, it is not that diverse in terms of where these markets are hosted. NYSE/ICE host the majority of these markets in their Mahwah datacentre, with Nasdaq hosting the majority of their markets in Carteret. The Equinix Secaucus site hosts many other exchanges and trading in different assets.

This means that collecting and monitoring the huge amount of data generated at these sites can be a challenge for a firm looking to collect that data and train models on it.

Scalable Capture

First of all, Beeks Analytics has the scalable capture to be able to store this information efficiently. Our latest generation of appliances can capture data at 200Gbps and, thanks to our FPGA compression technology, can store up to 2 petabytes of packet captures on the appliance.

If two petabytes aren't sufficient for you, then our software allows you to offload packet captures to object-based cloud storage, whether for archiving, batch-based analysis, or regulatory purposes. Our data archiving software splits packet captures by market and provides cloud sync, saving you processing and development time. These cloud-based stores can act as the seed beds for future AI analysis work, if you are still developing your approaches. By storing every packet - sensibly categorised by market - firms gain unparalleled historical visibility enabling enhanced AI model training.

Flexible Decoders

However, once your central AI initiative starts to process data you've stored in the cloud or on your packet capture appliances, you'll need to make sense of the packet data. Large Language Models show promise in analysing PCAPs one-by-one, but to use PCAPs for AI you need to systematically extract features from them rather than feeding them into a large model.

The Beeks Analytics VMX-Capture component doesn't just provide the software and qualified appliance hardware for writing the captures. It also provides the feature extraction which AI models demand.

Our VMX-Capture network decodes are much more efficient at bulk PCAP processing than free tools like Wireshark, tshark, pyshark and their associated tools. And with

200+ financial markets and enterprise decoders, you don't need to waste valuable development resources building and maintaining these yourself.

These decoders can both output statistical summaries of the packet capture contents and provide normalised message output of some or all of the packet streams. These normalised outputs will be easier to work with for many AI use cases than the raw packet captures. The flexible nature of the VMX-Capture configuration allows you to extract the fields you need from the packet captures to simplify your AI workload.

Streaming Use Cases

However, there are other use cases where you need streaming access to wire data. Perhaps you need the data as close to real-time as possible. Or perhaps you need systematic collection of session information about the conversations that are taking place over your network.

We can provide statistics and telemetry over Kafka (see page 32) or as an Open Telemetry Publisher.

Streaming is also a good technique where you don't want the full packet capture – maybe you only want to look at certain fields from the messages.

Our high performance decoding layer can capture, decode and egress up to 300 million messages per second per appliance– giving you the visibility you need in your data lakes for some of the most demanding financial data.

For even higher volume requirements, we can stream using Apache Arrow or transfer the data using the Parquet data format, both of which allow you to reach even higher message rates.

Case Study Conclusions

In a market where the volume and velocity of financial data continue to rise, firms need a robust solution to capture, decode, and analyse this information efficiently. Beeks Analytics provides a comprehensive suite of tools that not only handle the scale of market data but also offer the flexibility needed for AI-driven insights. Whether through high-speed packet capture, advanced decoding, or real-time streaming, our solutions empower firms to extract value from their centralised data with minimal overhead. By leveraging scalable storage, efficient processing, and seamless integration with AI workflows, Beeks Analytics helps firms stay ahead in a competitive landscape where data-driven decision-making is critical.

What is Edge Intelligence?

Edge Intelligence (sometimes referred to as Edge Artificial Intelligence) refers to processes in which data is collected, analysed, and insights delivered close to where it is captured in a network. This technological approach represents the convergence of edge computing infrastructure with sophisticated artificial intelligence and machine learning capabilities. Unlike traditional cloud-based AI systems, which require data to be transmitted to centralised servers for processing, Edge Intelligence is a general industry approach which enables computing resources to be distributed throughout the network, particularly at its periphery where data originates.

This decentralisation is affecting multiple industries, from irregular heartbeat detection in medical wearables to wind turbines that collect vibration and temperature data to predict impending maintenance needs. The advances are now seen in the mass consumer market, with smartphones moving from Neural Engines that focus on solving specialised tasks like face recognition to chips that can run a much wider set of use cases, including Large Language Models.

The benefits of intelligence at the Edge vary per industry, but broadly include:

- Privacy - Sensitive data can be processed locally, reducing the need to transmit it to central servers or external datacentres.
- Latency - Processing data at the point of collection significantly reduces response time by eliminating the round-trip delay to a central system.
- Stability - the services that intelligence unlocks no longer need to rely on a continual central connection, and survive interruptions in network connectivity.
- Scalability - By distributing the processing across edge nodes, you can avoid the need to scale a central server.
- Contextual Intelligence - Incorporates the much richer data available local conditions (e.g., temperature, location) in decisions, improving relevance and accuracy.

In its broadest sense, Edge Intelligence can be conceptualised as a decentralised edge computing paradigm that brings high-performance computing capabilities to the richer contextual data available at the periphery of networks.

This arrangement allows systems to filter, process, and analyse information before determining what needs to be transmitted to centralised systems. Edge devices typically exchange processed insights, anomaly flags, or aggregated summaries with central systems, rather than transmitting every raw data point. This reduces bandwidth usage and latency, but it also makes centralised oversight more complex as organisations must manage and visualise data from the many of heterogeneous edge devices. Orchestration, automation, and flexible central analytics dashboards become crucial to managing the Edge Intelligence world, in all industries.

The market research firm IDC estimate the current global market for edge computing will grow from $228 billion in 2024 to $378 billion by 2028. This continued growth means continued technological innovation in areas such as hardware miniaturisation, energy-efficient computing architectures, and AI algorithms optimised for edge deployment.

Organisations that successfully navigate this evolving landscape will be well-positioned to realise significant competitive advantages through faster decision-making, reduced operational costs, and enhanced ability to derive actionable insights from the ever-increasing volumes of data generated at the network edge.

Edge Intelligence for Capital Markets

The development of the financial markets has been powered by the twin engines of global interconnectivity and physical proximity. The early modern financial centres often developed at or near busy ports, where merchants gathered to exchange goods and information. The traders of the 17th and 18th century relied on physical proximity to the latest shipping news, which helped traders and insurers quickly adjust prices and policies, creating an efficient marketplace.

As markets became increasingly electronic in the 21st century, the koffiehuizen of Amsterdam or of the Royal Exchange found their modern equivalents in the datacentres of New Jersey and Slough. As the move to digital platforms increased the velocity of trading, financial firms moved their servers closer (or inside) the exchanges to reduce latency. Neutral datacentre colo sites like those belonging to Equinix, Interxion, and Digital Realty prospered as they attracted financial customers and markets who wanted proximity to each other and to the smaller (often OTC or ATS/MTF) trading venues that equally saw a benefit in being close to their potential customers.

Because of this landscape, almost all financial firms with trading operations will have to trade across multiple venues in different physical locations. So **the Edge** is now well defined for financial markets, as being the locations at which trading connections are richest and within which trading is conducted. This contrasts with the core of financial firm's processing for less latency sensitive workflows (e.g. retail customers or settlements), much of which is moving to hyperscale cloud locations as vertically integrated financial firms face pressure for financial and energy efficiency.

Beeks Analytics and similar tools such as Corvil and Packets2Disk have grown up to help firms manage their latency-sensitive workflows. They provide real-time network and trade flow monitoring for low-latency, high-performance trading environments. These tools are essential for trading firms operating in colocation (colo) facilities, where every microsecond matters (for some firms, every picosecond).

With Market Edge Intelligence, Beeks brings this performance monitoring to its next stage by taking advantage of innovative technology to allow advanced AI-driven analytics at the edge, enabling real-time anomaly detection, predictive insights, and proactive decision-making for ultra-low-latency trading.

> *"Having smaller models which can operate entirely at the Edge (which for financial firms is often a colo or Exchange data centre) will help address latency concerns, as well as cost of use and privacy issues. Just look at what Apple, Samsung and Google are doing to optimise AI models on the upcoming generation of smartphones, and imagine the power of this applied to financial*

markets use cases."
Steve Rodgers, CTO in Analytics

WHY MONITORING AT EDGE IS SO IMPORTANT FOR FINANCIAL MARKETS

Financial markets operate at microsecond timescales, where physical distance between systems directly impacts trading performance.

Centralised monitoring architectures introduce unavoidable delays due to:

1. **Propagation latency**: Each kilometer of fiber adds ~5 µs delay (compounded by router processing)
2. **Serialization/deserialization overhead**: Converting binary market data formats (e.g., ITCH, FIXP) for long-haul transmission
3. **Queueing delays**: Congestion at centralised aggregation points during volatility spikes

What sets capital markets apart from other sectors is the intersection of time sensitivity, regulatory scrutiny, and data richness. Markets generate vast volumes of tick data, order flow, and network telemetry every second. By embedding AI capabilities directly into edge appliances in colocation facilities, firms can:

- **Pre-empt network issues** before they impact execution quality.
- **React to market microstructure changes** in real time, by spotting latency spikes, dropped packets, or unusual message patterns as they happen.
- **Drive execution improvements** through high-fidelity signal detection from raw wire data.

Beeks Market Edge Intelligence

Beeks has long recognised the huge opportunities and significant challenges of AI-powered analytics at the edge. We've launched Beeks Market Edge Intelligence, our edge AL/ML solution that delivers low-latency, high-context insights that give you better value from the data going across your network.

Beeks Market Edge Intelligence:

- provides low-latency forecasting, correlation, anomaly detection, similarity analysis, and root cause analysis for non-statistical users.
- offers direct access to GPU-accelerated hosted models and high-volume message stores.
- enables power users to experiment with data at the edge by querying data directly on the appliance, exploring patterns, testing hypotheses, and applying AI/ML techniques. Use a range of tools/interfaces, bring your own models to the edge data, and bring your own data to the Beeks platform. With our AI/ML-enabled SQL interface, Jupyter Notebooks, and VMX-Explorer Grafana dashboards, working with wire data at scale has never been easier.

And with our open architecture, it's easy to move your analysis from the edge to the centre, or vice versa.

Beeks Market Edge Intelligence can be implemented alongside Beeks Analytics to provide niche edge analytics and centralised broader analytics and replay.

MACHINE LEARNING APPROACH TO EDGE INTELLIGENCE

The ability to learn from the vast volume of captured wire data is what gives AI/ML such tremendous transformative power in the area of edge intelligence and predictive analytics. Machine Learning algorithms monitor performance and trading from the captured wire data, to learn patterns that are in the normal range (baseline) and identify anomalies that deviant from said norm. Edge Intelligence's advanced anomaly detection algorithms can further consider contextual factors like network environment, traffic types and routing, dataset variables, in order to determine how these factors interact to impact performance.

Forecasting and Predictive Maintenance

Edge Intelligence leverages historical performance data and live captured data to identify future performance risks to infrastructure like bottlenecks or failures. Combining multiple layers of information (e.g. loss, latency, TCP window sizing)

captured from the network unlocks predictive maintenance, where components that are likely to fail or introduce latency are taken out of service and replaced before they can impact performance thereby reducing operational risk.

HFT and quant firms are laser-focused on latency reduction and performance optimisation, as even microsecond delays can impact profitability. Traditional monitoring tools do not provide the predictive insights needed for ultra-low-latency environments. This forces operations teams to be reactive. An ML-powered solution in colo, at the Edge will leverage time-series data to pinpoint latency spikes, identify patterns in infrastructure performance, and suggest optimisations that reactive monitoring might miss. The ML can continuously analyse and recommend changes, such as network routing adjustments or hardware reconfiguration, to minimise latency dynamically.

> *"Being able to predict the future is gold in Capital Markets businesses. We can see that advancements in the Large Language Models (LLMs) underpinning GenAI are augmenting traditional Machine Learning (ML) techniques to allow for accurate forecasting with reduced costs and time to market."*
> Steve Rodgers, CTO Analytics

Capacity Planning and Scalability Forecasting

Trading volumes and data throughput requirements can vary substantially, particularly during market events or periods of increased volatility. Predicting future infrastructure needs is challenging, especially in environments where over-provisioning is costly. AI-based time-series analysis can predict usage trends, helping firms optimise their capacity planning for compute, storage, and networking resources. This would allow firms to provision resources dynamically, avoiding the cost of over-provisioning while ensuring scalability. One example can be market data line saturation. It would be prudent to proactively know when market data from a particular feed, or exchange can burst and cause significant quality issues. It would be prudent to have a solution which can recommend line upgrades given a sustained breach in thresholds.

Historical Trend Analysis for Post-Event Diagnostics

When infrastructure issues occur, post-event diagnostics are crucial for root cause analysis and future prevention. Manually identifying causative patterns in extensive time-series data is time-consuming and often inconclusive. AI can quickly scan historical data to identify patterns that correlate with past failures, providing actionable insights for preventing repeat issues. A "post-event analytics" tool can automatically generate detailed reports on infrastructure incidents, linking patterns to probable causes and suggesting preventive measures.

Similarity Analysis

AI-driven **similarity analysis** measures how closely related two data streams or events are, helping to reveal hidden correlations in trading network traffic. By extracting key features from real-time network telemetry and market feeds, similarity analysis can group together flows or behaviors that exhibit high correlation. This not only reduces data volumes (by focusing on representative streams) without losing critical information, but also highlights subtle relationships that manual monitoring might overlook. For example, if multiple price feeds or order streams carry redundant information, an AI at the edge can identify those highly similar streams so that traders and support teams can monitor a smaller subset while still covering all important data.

This technique also assists in anomaly detection: outliers become evident as data that **does not** fit any known similarity cluster, enabling quicker detection of abnormal network conditions or suspicious trading patterns.

Ultimately, AI-driven similarity analysis improves trading execution by ensuring that decisions and algorithms are based on comprehensive yet de-duplicated data. Traders gain clearer insight into network behavior and can trust that any anomalies (such as an exchange feed lagging behind others) will be promptly flagged, allowing them to react before it impacts trade performance.

Trading Signals

AI at the edge can extract and refine trading signals from raw network data in real time, providing a critical advantage in fast-moving markets. Instead of sending all data to central servers for analysis, edge AI models process price and order data alongside deep network health metrics locally to generate immediate insights. These insights range from predictive indicators (e.g. forecasting the next price move or liquidity shift) to alerts (e.g. detection of an arbitrage opportunity or an unusual surge in order flow). By filtering noise and focusing on high-impact patterns, AI transforms high-volume data feeds into actionable signals for trading algorithms. Performing this analysis at the edge provides three major benefits. Firstly, performing this analysis at the edge slashes response times – decisions can be made in microseconds at the exchange colocation, which is crucial for competitive automated trading. Secondly, undertaking this analysis at the edge allows the signals to be generated based on the rich wiretime accuracy and lower level network information which is lost when viewing information centrally or within the application layer. And thirdly, performing this analysis in parallel with your decision making systems minimises the amount of work that those systems have to do, allowing them to operate at reduced latency.

In practice, embedding AI alongside traditional analytics provides an additional *"steering signal"* for trading systems – guiding algorithms on when to execute or hold fire. Market participants leveraging edge-derived signals gain a tangible competitive advantage, as their trading decisions are informed by the most up-to-date, context-rich data possible. This real-time intelligence enables better timing of trades, more efficient pricing, and the ability to capitalise on fleeting market inefficiencies before others do, thereby materially improving execution quality and profitability.

ARCHITECTURE

Market Edge Intelligence takes the same open architectural approach as Beeks Analytics, making it easy to move your analysis from the edge to the centre, or vice versa.

Market Edge Intelligence introduces the CDF-Q interface. CDF-Q is a new data stream for Beek's Core Data Feed that provides the results of queries from the edge via Kafka in Market Edge Intelligence.

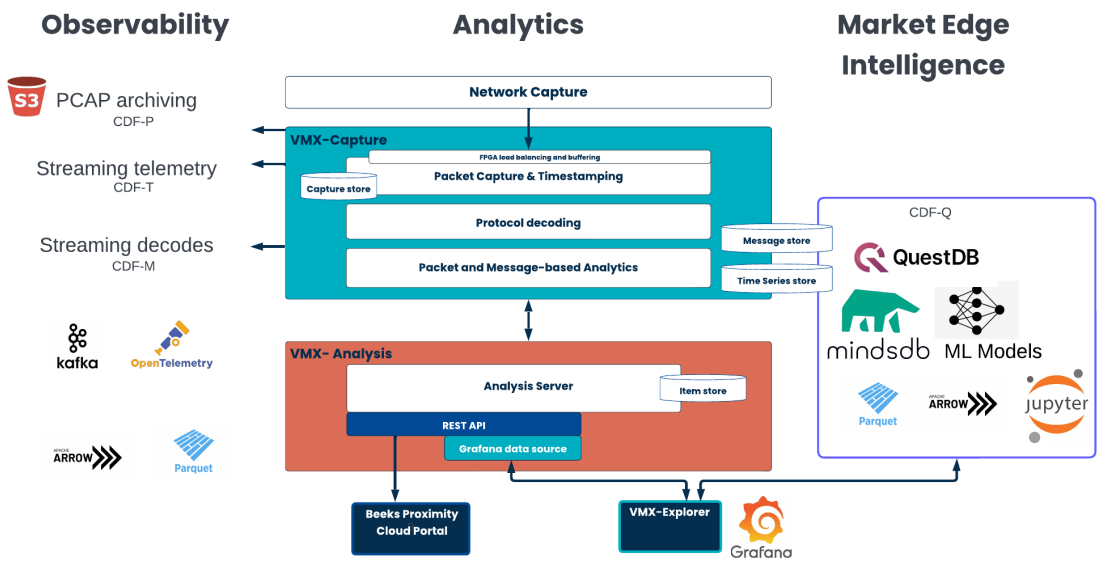

Figure 2 Edge Intelligence within Beeks Analytics Overview

Apache **Parquet** is an open-source column-oriented data storage format that supports large data sets because it's optimised for disk I/O and can achieve high compression ratios with columnar data. In Beeks Market Edge Intelligence, we can transfer high-volume data using the Parquet data format.

Apache **Arrow** is a streaming mechanism for big data that we can use in Beeks Market Edge Intelligence as an alternative to using Parquet format.

EDGE INTELLIGENCE COMPONENTS

Beeks Market Edge Intelligence is built on a modular, extensible foundation that ensures firms can scale their analytics efforts without sacrificing flexibility or performance. Each component of the solution has been designed to address a critical requirement of high-speed, AI-enabled analytics in trading environments—whether that's processing at the wire, visualising anomalies, or integrating your own models (or your own edge timeseries data) into the data pipeline.

Featurised Data

The core of Edge Intelligence is its access to rich, granular telemetry data from financial networks. This includes metrics from market data gateways, trading algorithms, and order execution paths. These data streams are "featurised" through Beeks' decoding layer, allowing users to extract precise, structured signals from raw packet flows. The resulting data is optimised for time-series analysis, anomaly detection, and root cause investigation.

Exogenous Metadata

To enhance model accuracy and contextual awareness, Market Edge Intelligence incorporates external metadata such as economic calendars, trading hours, weather data, and regulatory holidays. These inputs enable correlation with external events, improving forecasting models and supporting more intelligent resource allocation (e.g., scaling gateways in anticipation of high-impact announcements).

Time-Series Database

At the heart of the platform is QuestDB, an open-source, high-performance time-series database. QuestDB provides a columnar, SQL-native interface for ultra-fast querying and analytics, whether you're accessing wire telemetry or decoded trading messages. QuestDB provides real-time ingestion and historical queries, enabling low-latency dashboards and in-depth forensic analysis on the same infrastructure. The open nature of QuestDB makes combining Beeks Analytics and your own timeseries data easy, if you want to perform this close to the source of the data.

AI Middleware

The AI middleware layer orchestrates machine learning workflows by integrating with frameworks such as MindsDB, AutoTS, and Keras. This layer automates training and deployment pipelines, supports BYOM (Bring Your Own Model), and bridges high-speed data from QuestDB into actionable ML insights. Whether you're running

anomaly detection, clustering, or predictive maintenance models, the middleware layer ensures consistent model deployment and observability.

Scalable Compute

Edge Intelligence supports distributed AI execution through GPU-enabled infrastructure deployed directly in colocation environments. This makes it possible to train and execute complex models—such as deep learning forecasting models—at the edge, without needing to send all of the data to central locations.

Data Visualisation and Exploration

The decoded and modelled outputs from Edge Intelligence can be explored using Beeks' custom dashboards via VMX-Explorer (Grafana), or through hosted Jupyter Notebooks for more advanced, code-driven exploration. These visual interfaces make complex analytics accessible to traders, operational teams, and engineers alike—supporting everything from alert investigation to hypothesis testing.

The system provides full data transparency—whether you're extracting PCAPs for low-level inspection or visualising streaming metrics for infrastructure tuning.

Alerting

Edge Intelligence includes a flexible event notification system that can push alerts based on thresholds, anomalies, or custom-defined patterns. This is critical in high-frequency trading environments, where immediate awareness of performance degradation or unexpected latency can protect revenue and ensure compliance.

Conclusion

The future of financial analytics lies at the edge—where data is generated, and decisions must be made in microseconds and faster. Beeks Market Edge Intelligence empowers firms to meet this future head-on, delivering AI-powered insight directly into colocation facilities, trading gateways, and network perimeters.

With native integration of high-performance tools like QuestDB and Kafka, along with its already strong financial protocol decode and packet capture capabilities, Beeks enables a new generation of flexible, scalable, and vendor-neutral analytics workflows. This architecture gives firms full control over their wire data, breaking free from the constraints of proprietary appliances and siloed data systems.

Whether you're training machine learning models, optimising market data flows, or detecting anomalies before they become outages, Beeks Market Edge Intelligence equips you with the tools to act faster, predict more accurately, and gain deeper insight—where it matters most.

For more information about Beeks Market Edge Intelligence, please contact sales@beeksgroup.com.

Appendix: Why Beeks Analytics chose QuestDB as our database engine of choice to unlock the value of your wire data

Beeks Analytics and before has had a long-standing commitment to openness, value and avoiding expensive vendor lock-in.

This continues with our implementation of QuestDB for our open access timeseries and message store. Beeks provides access to this database using our CDF-Q interfaces.

By contrast, Corvil data has to be exfiltrated via Analytics streams (limited to 10 per appliance) or viewed in an expensive and obsolete Hadoop-based 'iHub' central store (which provides a UI for queries but no open data access).

Storing this data in QuestDB provides the following advantages:

/ SQL Support

QuestDB offers a standard SQL interface for querying time series data. This is easier to adopt compared to proprietary appliances that may only provide a limited or custom API.

/ Advanced Time Series Extensions

QuestDB supports time series–specific SQL extensions (e.g., time-based aggregations, sampling, downsampling, time-partitioned tables), making it straightforward to analyze large volumes of data.

/ Columnar Storage & Vectorized Execution

By storing data in a columnar format and using vectorized query execution, QuestDB can achieve very low-latency queries over huge datasets—crucial for real-time monitoring or analytics.

/ Parallel Execution & In-Memory Capabilities

QuestDB takes advantage of modern CPU architectures, parallelizing operations to handle large concurrency and data volumes more efficiently than many proprietary systems that might be constrained by their fixed hardware design.

/ Rich ecosystem of connectors and integrations

Access your data in QuestDB using SQL, JDBC etc, providing easy integration with your own Business Intelligence tools. Or use Beeks Market Edge Intelligence to integrate the data into your AI/ML workflows.

/ Full Data Ownership

With a proprietary appliance offering limited API endpoints, your ability to explore and manipulate data might be constrained by what the appliance vendor allows. QuestDB, on the other hand, lets you query data directly and in more complex ways, ensuring you can derive any insights you need.

/ Future-Proofing & Portability

If your requirements evolve or you want to switch to a different technology stack, it's simpler to migrate data from an open-source system than from a closed appliance with proprietary data formats and APIs.

IN-DEPTH REVIEW OF QUESTDB'S BENEFITS FOR OPEN DATA ARCHITECTURES

QuestDB stands out in the time-series database landscape by offering superior integration and extensibility compared to proprietary data appliances. QuestDB's architecture provides significantly more flexible data access patterns, interoperability with modern data ecosystems, and cost-effective scaling compared to proprietary alternatives that typically restrict data access through limited APIs and create deliberate vendor lock-in.

The Integration Limitations of Proprietary Data Appliances

Proprietary data appliances have long presented challenges for organisations seeking to build integrated, flexible data architectures. These closed systems typically restrict access to data through tightly controlled, often limited APIs that serve the vendor's business interests rather than the customer's integration needs. Many proprietary time-series appliances force organisations into predetermined workflows, making it difficult to extract, transform, and analyze data across multiple systems. This controlled access creates artificial barriers between the data storage layer and applications that need to consume that data, resulting in slower development cycles and reduced analytical capabilities.

The financial implications of these limitations are significant. Proprietary vendors typically charge premium prices for additional connectors, integration modules, or expanded API access that should be standard. This tiered access model means organisations often pay repeatedly for access to their own data across different systems. Furthermore, these systems frequently store data in proprietary formats that make migration to alternative platforms technically challenging and expensive, creating a form of technical debt that compounds over time as data volumes grow.

Scaling proprietary appliances presents another set of challenges. Most proprietary solutions require purchasing additional hardware or licenses from the same vendor, preventing organisations from leveraging commodity hardware or open-source technologies to manage growth cost-effectively. This vendor lock-in extends to the entire data pipeline, as proprietary systems rarely offer seamless integration with the broader data ecosystem without additional costs or compromises in functionality or performance.

QuestDB's Open Architecture and Integration Philosophy

QuestDB approaches time-series data management with fundamentally different design principles that prioritise integration and extensibility. As an open-source time-series database designed for high-performance analytics, QuestDB employs a columnar storage model that naturally aligns with modern analytical workloads. Unlike proprietary alternatives that limit access points, QuestDB provides multiple ingestion and query interfaces including PostgreSQL wire protocol, InfluxDB line protocol, REST API, and direct file imports. This multi-protocol approach ensures organisations can connect existing tools and applications to QuestDB without expensive middleware or customisation.

The database's design philosophy centers on openness and interoperability rather than creating dependencies. QuestDB's core storage engine is optimised for time-series data while maintaining compatibility with SQL, the universal language of data manipulation. This SQL compatibility, enhanced with specialised time-series extensions, allows analysts and engineers to leverage existing skills and tools rather than learning proprietary query languages or interfaces. The open architecture extends to deployment flexibility, with QuestDB supporting everything from embedded applications to large-scale distributed environments without forcing architectural decisions that benefit the vendor rather than the user.

QuestDB's commitment to open standards is particularly evident in its implementation of industry-standard protocols. By supporting the PostgreSQL wire protocol, QuestDB enables immediate integration with the vast ecosystem of PostgreSQL-compatible tools, languages, and frameworks. Similarly, support for the InfluxDB line protocol allows seamless migration from InfluxDB deployments and integration with tools designed for that ecosystem. This protocol diversity eliminates the need for proprietary connectors or gateways that add cost and complexity to data architectures built around closed systems.

Apache Parquet: Enhancing Data Portability and Analytical Performance

QuestDB's integration with Apache Parquet represents a significant advantage over proprietary systems that rely on closed data formats. Parquet is an open columnar storage format that provides efficient data compression and encoding schemes

specifically designed for analytical workloads. By supporting Parquet import and export, QuestDB enables seamless data exchange with the broader data ecosystem, including major platforms like Hadoop, Spark, and various cloud data warehouses. This interoperability is crucial for organisations that need to incorporate time-series data into larger analytical workflows without expensive data transformation or proprietary conversion tools.

The technical advantages of Parquet integration are substantial. Parquet's columnar format stores data by column rather than by row, which dramatically improves performance for analytical queries that typically access a subset of columns. This alignment with analytical access patterns means queries run faster while consuming fewer computing resources. Additionally, Parquet's sophisticated encoding and compression schemes reduce storage requirements and I/O operations, further enhancing performance and reducing costs compared to proprietary formats that often prioritise vendor lock-in over efficiency.

Parquet's schema evolution capabilities provide another layer of extensibility that proprietary formats rarely match. As data requirements evolve, Parquet allows for schema changes without breaking existing queries or requiring data migration. This flexibility is particularly valuable for time-series data, where new metrics or dimensions might be added over time. QuestDB leverages these capabilities to provide a future-proof foundation for evolving data needs, whereas proprietary systems often require expensive data conversion or migration projects to accommodate changing requirements.

The integration between QuestDB and Parquet extends beyond basic import/export functionality. QuestDB's native columnar storage model aligns naturally with Parquet's design, enabling efficient data movement between storage and memory without the performance penalties associated with format conversion in systems not designed with columnar principles. This architectural harmony allows QuestDB to maintain high performance across diverse workflows that might involve importing data from Parquet files, performing time-series analysis, and then exporting results back to Parquet for consumption by other systems in the data pipeline.

Apache Arrow: Revolutionizing In-Memory Analytics and System Interoperability

QuestDB's implementation of Apache Arrow represents perhaps its most significant technical advantage for integration and extensibility. Apache Arrow provides a standardised columnar in-memory format and a set of libraries for efficiently moving data between different systems without serialization overhead. By supporting Arrow, QuestDB enables true zero-copy data sharing with other Arrow-compatible systems, eliminating the performance bottlenecks associated with data conversion and serialization in traditional database integration approaches.

The technical implications of Arrow integration are profound. When transferring data between systems that both support Arrow, the overhead of serialization and deserialization is eliminated entirely. Data stays in the same memory layout throughout the process, allowing systems to share data at memory bandwidth speeds rather than being limited by CPU-intensive conversion processes. This capability is particularly valuable for time-series analytics, where large volumes of data often need to move between specialised systems for different types of processing.

Arrow's columnar in-memory format aligns perfectly with modern CPU architectures, enabling vectorized operations that process multiple data points simultaneously. QuestDB leverages these capabilities to accelerate analytical queries beyond what's possible with row-based in-memory representations. This performance advantage extends across the entire data pipeline when Arrow is used consistently, enabling analytics at speeds that proprietary systems with non-standard memory formats cannot match without specialised hardware accelerators that further increase costs and vendor dependencies.

The integration between QuestDB and Arrow extends to language bindings and client libraries. Arrow provides interfaces for multiple programming languages including Python, R, Java, and C++, allowing developers to work with QuestDB data in their language of choice without performance penalties. This language flexibility stands in stark contrast to proprietary systems that often provide first-class support for a limited set of languages or force developers to use vendor-specific client libraries that may not align with modern development practices or organisational standards.

Practical Applications: Real-World Integration Scenarios

The combination of QuestDB with Parquet and Arrow creates integration possibilities that would be technically challenging or prohibitively expensive with proprietary alternatives. In financial services, for example, QuestDB can ingest market data streams in real time, perform complex time-series analysis, and then make the results available to risk management systems through Arrow interfaces at memory speeds. This high-performance data sharing enables real-time risk assessment that wouldn't be possible with the API bottlenecks typical of proprietary systems.

For machine learning applications, QuestDB's Arrow integration enables direct data transfer to frameworks like PyTorch or TensorFlow without the serialization and deserialization overhead that typically bottlenecks ML pipelines. Feature engineering can be performed in QuestDB using SQL, with the results flowing directly to model training processes at memory speeds. This streamlined workflow eliminates the data preparation bottlenecks that often consume the majority of time in ML projects when using databases with limited integration capabilities.

Cost and Performance Implications of Open Integration

The economic advantages of QuestDB's open integration approach extend beyond eliminating license fees for connectors or integration modules. By supporting open standards like Parquet and Arrow, QuestDB reduces the need for intermediate data transformation layers that add complexity and cost in architectures built around proprietary systems. Organisations can build direct data pipelines between systems that speak these common languages, eliminating both the licensing costs and computational overhead of proprietary middleware.

Performance metrics further illustrate the advantages of QuestDB's approach. When transferring data between Arrow-compatible systems, throughput improvements of 10-100x compared to JSON-based API transfers are common. These performance gains translate directly to reduced infrastructure costs, as the same analytical workloads can be completed with fewer compute resources. Similarly, Parquet's compression capabilities typically reduce storage requirements by 75% compared to row-based formats, creating significant cost savings for large-scale time-series data storage.

The total cost of ownership advantages become even more apparent when considering the entire data lifecycle. QuestDB's open design allows organisations to evolve their data architecture incrementally, adding or replacing components as needs change without forklift upgrades or data migration projects. This architectural flexibility eliminates the costly rip-and-replace cycles common with proprietary appliances, where changing requirements often necessitate complete system replacements rather than incremental evolution.

Conclusion: The Future of Time-Series Data Integration

QuestDB's implementation of open standards like Apache Parquet and Apache Arrow represents a fundamental architectural advantage over proprietary data appliances that limit access to data through restricted APIs. By embracing these open formats, QuestDB enables seamless integration with the broader data ecosystem, high-performance data sharing between systems, and flexible data architectures that evolve with organisational needs. This open approach stands in stark contrast to proprietary systems designed to create technical dependencies and ongoing revenue streams through artificial integration barriers.

The integration advantages of QuestDB extend beyond current capabilities, as the ecosystem around Parquet and Arrow continues to evolve. New tools and systems that support these standards automatically become integration points for QuestDB without requiring vendor-specific development or additional licensing costs. This network effect creates increasing value over time, as more components of the modern data stack adopt these open standards for efficient data exchange and processing.

For organisations building data architectures for long-term flexibility and performance, QuestDB's open integration approach offers clear advantages over proprietary alternatives. By eliminating artificial API barriers, embracing columnar formats like Parquet for storage and Arrow for memory representation, and supporting multiple access protocols, QuestDB provides a foundation for time-series data management that prioritises integration and extensibility rather than vendor lock-in. As data volumes continue to grow and analytical requirements become more complex, this architectural openness becomes increasingly valuable for organisations seeking to extract maximum value from their time-series data.

Appendix: Beeks Core Data Feed

The Core Data Feed is the Beeks interface for streaming data out of our appliances.

The Core Data Feed:

- Streams analytics data like network telemetry, network session information, network metrics, or business information like correlated business transactions.
- Can provide an alert stream.
- Stream decoded messages (captured from the wire) to downstream systems for tick storage, quant analysis, etc.
- Access data using the CNCF Open Telemetry framework – move beyond log data by integrating metrics, spans and distributed traces.

It can stream at extremely high volumes - at STAC in October 2024, we presented our ability to stream messages at 300 million messages per appliance. This makes it easy to keep your big data systems up-to-date with decoded data from the wire.

NATIVE KAFKA SUPPORT FOR ENTERPRISE DATA ENGINEERING

A key advantage of the Beeks Core Data Feed (CDF) lies in its **native integration with Apache Kafka**. Rather than being constrained by proprietary streaming protocols or appliance-specific connectors, the CDF can seamlessly publish analytics events and decoded message flows into a standard **Kafka** infrastructure. This approach offers several benefits for large-scale capital markets environments:

1. **Massive Scalability and Throughput**
   Kafka's partition-based architecture enables near-linear horizontal scaling. You can add brokers to handle higher message rates, ensuring consistent low-latency delivery even as data volumes surge. This is relevant, for example, to exchanges dealing with tens of millions of messages per second—or for firms pushing real-time analytics into downstream systems without buffering or lag.
2. **Enterprise-Grade Reliability and Durability**
   Built-in replication and fault tolerance mean that even if individual brokers fail, your data stream remains intact and highly available. This resilience is key in financial services, where lost or partial data can have regulatory and operational implications.
3. **Open Ecosystem for Data Engineering**
   With Kafka as the backbone, Beeks Analytics users can integrate with a vast

range of data engineering and analytics tools—such as Apache Spark, Flink, or traditional ETL pipelines. Because Kafka is already entrenched in many enterprise IT stacks, it's straightforward to plug the Core Data Feed into existing data lakes, BI dashboards, or real-time analytics frameworks. There's no need to rely on proprietary interfaces that limit interoperability or require specialised connectors.

4. Supports AI/ML Initiatives
   AI and machine learning workflows thrive on continuous streams of rich, high-volume data. By leveraging Kafka's scalable publish/subscribe model, you can feed wire-level messages, order events, or latency metrics directly into ML pipelines. Data scientists and quant developers gain immediate access to high-fidelity data, fueling iterative model training and real-time inference. This native Kafka approach helps your organisation keep pace with the data-hungry needs of modern AI strategies—without the bottlenecks often seen in closed, appliance-based solutions.
5. Lower Operational Overhead
   Adopting Kafka for streaming data means you can rely on well-established, widely documented best practices for deployment, monitoring, and maintenance—rather than juggling multiple vendor-specific or limited streaming tools. The open-source nature of Kafka also means you avoid extra licensing costs tied to ingestion, scaling, or replication.

Native Kafka support is more than just another interface on the Beeks platform; it is a strategic enabler that aligns with modern data engineering principles and AI-driven business priorities. By integrating the Beeks Core Data Feed directly with Kafka, you gain a flexible, robust, and future-proof mechanism for transporting mission-critical network and trading data into every corner of your enterprise ecosystem—unlocking insights faster and at greater scale.

CORE DATA FEED ACCELERATION WITH QUESTDB

Newer versions of Beeks Analytics introduce QuestDB as the fast, open time-series database at the core of the system, enabling efficient SQL-based querying over wire data. To enhance portability and high-speed data exchange, QuestDB integrates seamlessly with Apache Parquet and Apache Arrow.

Apache Parquet allows Beeks to export decoded wire data in a compressed, columnar format that's ideal for storage and analytical workflows—making it easy to move data into cloud data lakes or downstream tools like Spark, Snowflake, or Python-based analytics.

Apache Arrow, in contrast, is optimised for **in-memory analytics** and inter-process communication. It's used for **real-time streaming of decoded telemetry**, especially when ultra-high throughput and low latency are critical. Arrow enables zero-copy transfers between systems—eliminating bottlenecks common with JSON or CSV formats.

By combining QuestDB's SQL-native time-series capabilities with Parquet and Arrow, Beeks delivers both flexibility and speed: historical insights at scale, real-time streaming at the edge, and seamless interoperability with the broader data ecosystem. Together, they unlock data agility across the entire AI and data storage pipeline - from packet capture to predictive analytics.

KEY ADVANTAGES OF BEEKS CORE DATA FEED

The Beeks Core Data Feed provides decentralised capture and centralised processing, significantly enhancing scalability and flexibility.

Key advantages of Beeks' architecture:

- Decentralised Capture with Centralised Processing: Beeks' system captures data at a single point and transmits it to a centralised analytics server, reducing infrastructure overhead.
- Scalability: A single Core Data Feed can handle multiple data streams without requiring additional appliances or independent analytics streams, making it horizontally scalable.
- Customizable and Efficient Data Delivery: Beeks CDF allows users to configure what specific metrics or decoded messages they need, reducing noise and improving efficiency.
- Lower Risk Exposure: The centralised model reduces the need for continuous monitoring of multiple streams, lowering the risk of missing critical trading data due to a stream failure.

SUMMARY

Feature	Beeks Analytics
Architecture	Centralised Core Data Feed with decentralized capture.
Scalability	One Core Data Feed handles multiple streams, enabling horizontal scalability. Utilizes Kafka partitions for near-linear horizontal scaling to handle hundreds of millions of messages per second.
Operational Complexity	Lower – single data source with customizable feeds.

Integration & Ecosystem	Standard Kafka topics easily feed Spark, Flink, BI tools, Big Data and AI/ML pipelines.
Data Pipeline Flexibility	Highly configurable publish/subscribe model with user-defined topics and filters.
Time to Insight	Low-latency feed directly into analytics or AI via Kafka; minimal extra overhead.
Cost & Overhead	Commodity servers + open-source Kafka reduce TCO; seamless scale without re-buying appliances.
Failure Risk	Lower – centralized processing reduces dependency on multiple streams. Kafka replication ensures data integrity even if broker nodes fail. Centralized analytics with decentralized capture plus Kafka's fault tolerance reduce risk.

Table 1 Summary of features

By building the Core Data Feed around Kafka, Beeks Analytics not only simplifies high-volume data capture but also future-proofs your data pipeline. This open, robust approach eliminates the scaling bottlenecks and lock-in risks tied to legacy appliance-centric architectures, enabling you to meet the exacting demands of modern trading, compliance, and AI/ML initiatives.

Document Shortcuts

TABLE OF FIGURES

TABLE OF TABLES