

Solving the Extreme Analytics Challenge in Financial Services

Inside

- 1 / Introduction
- 2 / Data intensive use cases in Financial Services
- 3 / The challenges
- 4 / Building a technology-powered competitive advantage
- 5 / Solving for better decision making and increased productivity
- 6 / Kinetica – A Technology Overview
- 7 / Conclusion

1 / Introduction

Dealing with big data is nothing new for financial services organizations. Banks have long understood the value of their data and applied analytics and machine learning to gain an advantage. But as banks continue their digital transformation, they are gaining access to more data sources than ever before in the form of financial data, news, transactions, social media, geospatial information, video, and much more. This data can be structured, unstructured, static, streaming, long-lived or perishable.

The world has moved far beyond the initial definition of big data. All this new data is unpredictable, challenging to manage, and complex to analyze. It will only get harder in the coming years as many new data sources come online with the exponential growth of the Internet of Things (IoT). This new data paradigm is called extreme data and it requires new methods of analysis. The banks that are able to gain insight from extreme data will take the lead as legacy systems struggle to process its complexity.

In an industry where speed matters, the ability to analyze data in real-time is critical. With so much data available, banks need to be able to pinpoint what is important and discard what is not. An [Accenture survey](#) found that 84% of bankers increasingly use data to drive critical and automated decision-making. This means the speed, quality, and depth of analysis of extreme data plays a major role in banking success.

In the Extreme Data Economy, data no longer just informs the business, it powers it. Data is how banks are looking to protect their balance sheet, know their customer, and make money. Leading financial services organizations are looking to advanced analytics and AI platforms as the next solution to their data challenges, be it risk management, customer experience, or more profitable trading strategies.

As banks seek to consolidate and automate systems, shift the emphasis to digital offerings, maximize the impact of technology budgets, and move faster to deal with tech-savvy competitors, innovative new solutions are needed that are purpose built for extreme data. It is the actionable business insights that are derived from extreme data that make data one of the bank's most valuable assets. Modern data and analytics platforms that can address the Extreme Data Economy are a critical component of successful innovation strategies.

“
Information is simply
as valuable as money.”

Source: Don Callahan, Head of Operations and Technology, Citi in [Rewiring Citi for the Digital Age - McKinsey](#)

2 / Data-intensive use cases in Financial Services

Across various areas of the bank, using data to power the business is critical. For some business units and teams, data is a core asset for decision-making and for monetization, but given the explosion of data, as well as the real-time nature of the data streaming in, existing technologies are hitting their boundaries, particularly for data-intensive and compute-intensive analysis. Core use cases that need advanced analytical capabilities include:

Capital Markets

- **Risk management**

Real-time risk management is a major problem facing the financial industry today and it is pushing conventional computing to its limits. The traditional historical simulations approach used for calculation of exposure and risk has significant drawbacks and more sophisticated methods like Monte Carlo are increasingly being used for risk and exposure calculations. These, however, are extremely data and compute intensive and complex to operationalize.

Value at risk (VaR) This is the statistical technique most commonly used by investment and commercial banks to determine the extent and occurrence ratio of potential losses in their institutional portfolios. VaR modeling determines the potential for loss in the entity being assessed, as well as the probability of occurrence for the defined loss. The common methods used for VaR calculations make assumptions about stock returns, like history repeating itself or a normal distribution on returns. Monte Carlo methods, on the other hand, involve developing a model for future stock price returns and running multiple hypothetical trials through the model. This provides a more comprehensive view of risk, but is more compute intensive.

Counterparty credit risk (CCR) It is important for banks to have a good understanding of counterparty credit risk for a range of instruments, including credit default swaps, as these type of instruments are often traded directly with another party. Since the contract is directly with the other party, there is a greater risk of counterparty default, since both parties may not have full knowledge of the financial health of the other. Potential future exposure (PFE) calculation with CCR is increasingly moving toward Monte Carlo methods as well.

- **Regulatory compliance**

Even though compliance capabilities have improved dramatically over the years, rising costs and emerging risks continue to put pressure on banks to develop efficient ways to address a constantly changing and complex regulatory environment. Accenture's "2017 Compliance Risk Study" found that nearly nine out of 10 (89%) compliance executives at banks, insurers and capital markets firms expect their costs to increase in their compliance department over the next two years. With an explosion in digital data and a finite pool of compliance talent from which to draw, firms are scrambling to find needed resources while keeping costs in line. Reporting and analytics are now an integral part of the Basel III regulatory accord. However, the use of legacy databases for calculations of key metrics and ratios including liquidity coverage ratio (LCR) means results take hours or days to be computed instead of minutes.

2 /

Retail

- **Customer 360 for omni-channel experience**

Banks generate massive amounts of data about their customers, from credit card transactions, bank visits, online banking interactions, and engagements across other channels. The explosion of the IoT creates a new channel of real-time streaming data that means banks are now continuously connected to their customer. Whether in-branch, on the phone, in the car, or interacting via another channel, banks can now engage their customers anytime, anywhere, on the channel of their choice. Data analytics is at the heart of delivering this 360-degree, personalized, omnichannel experience that delights customers and builds brand loyalty. With more ways to know the customer than ever before, traditional relational databases can't keep up with the volume, variety, velocity, and most of all, unpredictability of customer data.

- **Fraud detection and creditworthiness**

As the methods of digital payment increase, the methods by which credit fraud could occur increase as well. Multiple sources of data from all channels need to be aggregated in order to gain a holistic view of a customer's activity and behavior. Location data captured based on physical location of devices is growing in importance, as geospatial visualization can help to improve fraud detection. Additionally, the ability to analyze various types of data can enable banks to make better credit decisions, faster.

3 / The challenges

Siloed Systems

As banks grow organically and through mergers and acquisitions, their data infrastructure inevitably expands. Over time, this can lead to a sprawl of siloed systems, which makes it difficult to analyze data across business units and organizations. As a result, banks face challenges including mispriced assets, suboptimal operational efficiency, and an incomplete view of business critical elements like risk. It is also costly and inefficient to maintain redundant systems and report across them. The additional challenge of dealing with legacy systems that are unable to scale to meet new data requirements means banks are searching for a new solution that can help to connect the dots between disparate data sources and systems. As a full “heart transplant” to replace existing systems of records is unlikely, banks need to find a standards-based data platform that can help to ease the transition and bridge the gap between legacy and modern services. The [Boston Consulting Group](#) recommends banks have a data layer that is capable of connecting data clusters across the bank.

Legacy Technology Limits Bank Effectiveness

Over the years, banks have built data infrastructure to solve the challenges of the present day. However, these existing processes and legacy technology are not optimized for the new world of extreme data. What may have worked for big data struggles to keep up with the added complexity and unpredictability of extreme data. Legacy relational databases are performance challenged, leading to slower time to insight and undermining bank operations that need to be performed in real-time like risk management and trade decisioning. Agile business decisions are hamstrung by slow data analysis. Alternately, other financial services optimized solutions like KDB+ require specialist employees with a particular language skill set (Q language) and can be difficult to manage. They lack the ability to integrate machine learning models and also don't support a wide variety of use cases including location intelligence or counter-party risk. Unable to address challenges of today and tomorrow, legacy solutions are inadequate to deal with extreme data.

Inability to Analyze Extreme Data

While extreme data offers banks great opportunities to leverage new data points to improve their view of operations, it also brings new challenges. With multiple kinds of data coming from several different sources, workloads are becoming too large, unpredictable, and complex for standard relational databases leading to slower response times. Banks require the ability to aggregate data and find relationships across it in real-time. Business critical tasks like advanced risk measurements to appropriately measure credit and market risk for regulations, understanding customer behavioral patterns, and unearthing opportunities for cross-sell and upsell of financial products are all time sensitive. Similarly, analytics-based functions like credit decisioning, investments, underwriting, and stress testing call for both new kinds of data to further inform decision-making as well as speed of response.

On the customer-facing front, advanced segmentation and personalization for targeted marketing will be driven by data from all channels. This requires a new data platform that can not only connect data across siloed systems but also handle and quickly analyze new kinds of data.

Cybersecurity Threats

Banks have a big target on their back evidenced by the fact that the rate of breaches in the financial services industry has tripled over the past five years (Accenture). As custodians of not only financial assets but personal financial data, banks face significant cyber defense challenges. They must offer a superior level of security to maintain trusted customer relationships and risk losing clients if they cannot protect against malicious actors. Monitoring and understanding network data continues to get harder as more users, devices, and things become a part of the bank's digital ecosystem. Spotting and stopping unusual activity as it happens is crucial to maintain secure operations.

4 / Building a technology-powered competitive advantage

Key technology requirements for the data-powered bank

To increase agility and to enable faster digital transformation, it is imperative that next-generation data analysis and management technologies meet the key requirements that help solve the challenges faced by banks today. These requirements include

- **High-performance querying for data & compute-intensive analytics**

Support for high-speed online analytical processing (OLAP) including ad-hoc querying across very large data sets so that analysts can look beyond just a snapshot of the data and look into total exposure to the desired degree.

- **Unified platform to consolidate database sprawl**

Support for a large number of use cases both operational and analytical, resulting in consolidation of legacy relational databases like Oracle Exadata, columnar analytical databases like Vertica, and Cassandra, as well as some use cases of specialized databases like time series analysis.

- **Support for the most widely languages for a low learning curve**

Support for SQL, all the popular programming languages, modeling languages like Python as well as with RESTful APIs for custom application development. This helps firms reduce personnel costs as dependence on specialized skills decreases.

- **Real-time analysis on static and streaming data**

As the occurrence and importance of fast data increases, banks need next-gen technologies to not only ingest large amounts of data at high-speeds, but also simultaneously analyze and process the data to produce insights five to fifty times faster than existing systems.

- **Native support for multiple data types including location and time data**

Location and time data are not second-class citizens and can no longer be ignored. With the importance of this data increasing, data technologies need to support location-based and time-series analysis out of the box with a range of advanced functions.

- **Ability to bring algorithms and models to the data engine to operationalize all financial calculations and make them more efficient**

Running algorithms and calculations in custom-coded engines using Java or Python on CSV or Excel datasets is extremely error-prone. With the complexity of the algorithms and simulation techniques increasing, as well as the size of the data sets constantly growing, a next-generation engine will need to streamline processes and allow for machine learning and deep learning models to be run in-database.

Big data has enabled banks to turn into data-informed organizations. But as the types of data and the different data sources increase in the Extreme Data Economy, tools need to adjust as well. As banks continue their digital transformation, they need to not only rethink their data software stack, but the data analysis and processing software technologies need to fundamentally change the hardware foundation as well. GPU-accelerated databases can help financial services organizations with challenges that come with the unpredictability and complexity of extreme data. Kinetica offers an insight engine, built from the ground-up on a next generation hardware stack with a GPU database at its core to address all the key requirements needed to build a data-powered bank.

5 / Solving for better decision making and increased productivity

Banks are gifted with a wealth of data. So much so that the **Boston Consulting Group** believes banks should treat “data-driven intellectual property and analytics as one of their most valuable assets”. High-frequency trading firms, traditional asset managers, and traditional lending institutions can use Kinetica to measure risk, spot customer behavioral patterns, discover upsell opportunities and more. Kinetica’s GPU-accelerated database creates new opportunities to use data to gain a competitive advantage.

Capital Markets

Risk Management

Financial institutions can use Kinetica to perform risk calculations on-demand using the most up-to-the-moment data with sub-second speed. This allows banks to make better informed investment decisions and react quickly to market events, while reducing credit risk.

Kinetica is an ideal solution for complex risk calculations, as it removes the bottlenecks of data processing that arise with CPUs by leveraging the thousands of cores of general purpose GPUs. Parallel processing is a good fit for running risk queries multiple times a day to meet critical SLAs, as it can scale and deliver results in real time. Kinetica can also be used to accelerate complex market, liquidity or counter party risk calculations like liquidity coverage ratios, Monte Carlo simulations, xVA, and trade decisioning to obtain results intra-day rather than overnight. Kinetica can ingest data from any source and return instant insights, processing and analyzing billions of rows of data in milliseconds. Kinetica runs in-database analytics to handle time-sensitive, compute-intensive risk analysis at scale to project years into the future across hundreds of variables.

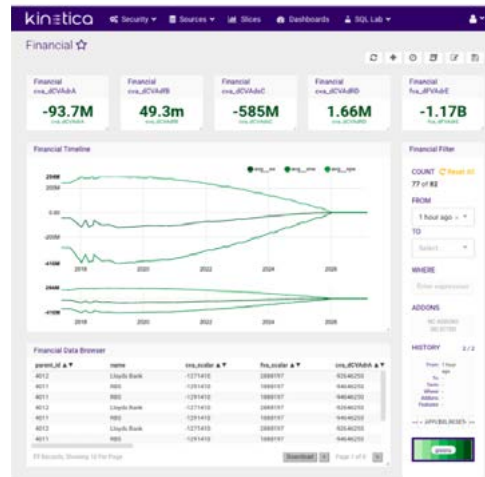


Figure 1 : Value at Risk - Credit value adjustments

Streaming Data

Real-time streaming analytics are increasingly important to the business as they enable banks to be more agile in decisioning, trading, and operations. With Kinetica, banks can continuously collect, analyze, and integrate streaming data with historical data for a more complete model. The speed, variety, and quality of information gives banks deeper insight into risk exposure, enabling rapid position adjustments to reduce risk. Analysts can run already developed code (Python, Java, etc) using UDFs within the database co-located with the data. In an industry where milliseconds matter and where insight directly equates to money, opportunities to apply machine learning and faster analytics offer a distinct competitive advantage. Kinetica makes it possible for financial organizations to gain knowledge and make predictions from vast volumes of complex and streaming data in milliseconds.

Trading

For trading, Kinetica can be used for tick data storage and analysis to deliver instant insight using time series analysis (time window analysis) for asset pricing. From there, investment professionals and traders can make more informed decisions at the speed of the market, not delayed by processing time. Analysts can model risk by bringing together many sources of trading data including instrument IDs, product codes, metadata, values, and more. Kinetica helps analysts to understand in real-time the failures of their models and where adjustments are needed, and provides them with greater insight into the data and a broader understanding of risk exposure. Increased transparency and analysis of risk exposure drives better business outcomes.

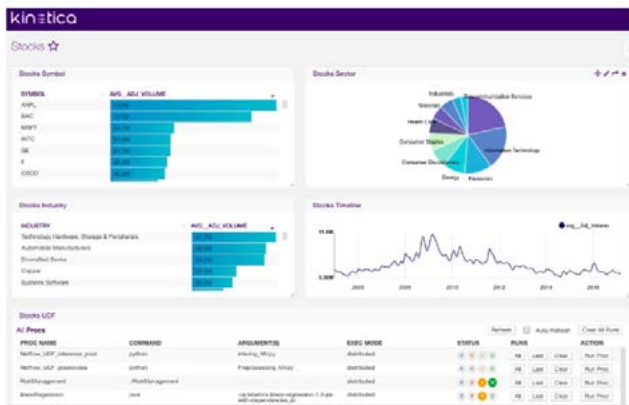


Figure 2: Ad-hoc analysis on large datasets with Kinetica

Retail Banking

Customer Experience

Kinetica can help to break down silos to connect transaction history with marketing and promotions by offering a single platform for data analysis. This helps banks to get a single view of the customer and their relationships across each channel. Financial services organizations can then use these insights to drive targeted marketing of loans and credit cards to select individuals, offer discounts based on geolocation or commute habits, identify high value customers, and even look ahead to the future with predictive capabilities as well as cross-selling and upselling. By analyzing large volumes of data from many different sources in real-time, banks can gain a granular understanding of customer behavior that allows them to interact with context and create relevant, valuable experiences for their customers. With Kinetica, banks can correlate data across spending habits, location, lifestyle, social media, connected devices, and more to build a more accurate picture of their customer. This can happen in seconds, not hours, making banks real-time businesses.

Operationally, using Kinetica, analysts can auto identify customer accounts at risk and execute credit line reductions or eliminations to reduce balance sheet charge-offs and loan loss reserves. This is critical to meet stringent reporting and compliance requirements. Data can even be used to address customer churn by identifying customers that are likely to leave and deliver targeted retention offers to rebuild brand loyalty. Now banks can find ways to turn customer data points into valuable information.

5 /

Fraud

Protecting against fraud goes hand in hand with customer experience. Kinetica can be applied for credit card fraud detection by combining geospatial data with transaction information and other customer data in real-time. Kinetica's visualization capabilities can be used for geo-fencing to identify fraud across regions particularly for mobile transactions. Additionally, fraud detection analysts can perform queries on large streaming datasets in order to uncover anomalies and patterns of behavior that signal potential fraud. By looking holistically at a customer's activity and applying advanced analytics and machine learning, banks can understand when behavior deviates from the baseline and flag it for review.

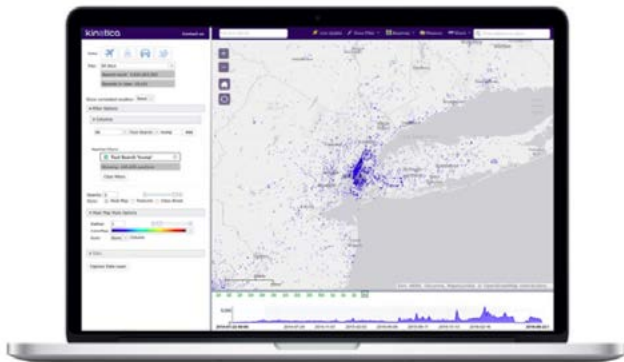


Figure 3 : Location-based fraud detection

Corporate / Across the Bank

Cybersecurity

GPU-accelerated data analytics can help banks to improve their cybersecurity posture. Kinetica can rapidly visualize and simulate multiple scenarios and reveal risk exposures, so that suspicious activity can be detected in seconds, not hours. Specifically, Kinetica can be used to combine data feeds with anomaly detection, monitor multiple streams of global attack vectors, find security lapses, and mine system logs. By applying machine learning and analytics, financial institutions can crunch petabytes worth of data in order to detect advanced persistent threats in real time and react quickly to abnormal or malicious activity

Reporting and Compliance

Driven by regulations like Basel II/III, banks are faced with significant reporting requirements, stress tests, and mandates to strengthen capital requirements, increase liquidity, and decrease leverage. Kinetica can help to move compliance report processing times from many minutes to seconds. As bank operations expand and become more complex with new financial products and exposure to different assets, solutions to model market, credit, and liquidity risk need to deliver results in seconds.



Figure 4 : Reporting with business intelligence tools powered by Kinetica

6 / Kinetica – A Technology Overview

GPUs for accelerated parallel computing

With the arrival of NVIDIA's CUDA drivers, the graphics processing unit has now turned into a general-purpose processing unit. GPUs crunch large volumes of data faster and more efficiently than CPUs because they work in parallel, instead of in sequence. GPUs accelerate location-based and in-memory analytics, machine learning, and AI.

Kinetica's instant insight engine is powered by thousands of advanced GPU cores that bring unparalleled speed, streaming data analysis, visual foresight, streamlined machine learning, and a best-in-class partner innovation ecosystem to break through the old bottlenecks.

Core Design Principles

Kinetica is built on three core design principles:

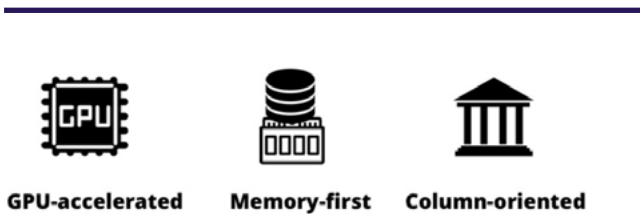


Figure 5 : Kinetica's Core Design Principles

- **GPU-accelerated**

The most fundamental architectural ingredients that make Kinetica unique are its in-memory GPU architecture and advanced GPU abstraction technology, that allow it to utilize the combination of CPUs and GPUs to deliver superior performance at a lower cost. This ingredient is special because most other similar systems are either not fully memory-based, or still use either GPU or CPU compute alone, leaving the other resource under or un-utilized.

- **Memory-first**

Kinetica adopted a memory-first architecture that includes using not only system memory or host memory but also chip memory or VRAM co-located with the GPU. This allows Kinetica to support not only quick operational point lookups, but also operational analytics with results in milliseconds. Data is also persisted and managed on disk.

- **Column-oriented**

Typically, analytical datasets are large because they contain huge amounts of historical data. Kinetica's native column-oriented design provides efficient data storage and blazing fast query performance for analytical datasets.



Figure 6 : Kinetica's Core Capabilities

6 /

Kinetica in your ecosystem

Interacting with Kinetica is no different than interacting with any traditional relational data platform, even though it has many architectural differences under the hood. Applications can send requests to any of the servers for processing via the SQL-92-compliant ODBC connector, or over native REST APIs that support JSON or avro serialization. This means that existing ETL and applications can stay, minimizing the time needed to get up and running with Kinetica.

With Kinetica, loading data into the cluster for analytics and reporting is designed to be simple and fast. Data records can be ingested into tables using distributed ingestion and extraction mechanisms. This makes it possible to load dimension and corresponding fact tables with billions of rows very quickly.



Figure 7 : Kinetica in your ecosystem

7 / Conclusion

The ability to extract insight from extreme data and transform it into a more valuable asset than the competition, in real-time, will be how banks find their edge. The insight engine with a GPU-accelerated database is the foundation for streaming data analysis, visual foresight, and streamlined machine learning that can help banks to be more nimble, productive, and informed. For both capital markets and retail banking, high-performance, advanced analytics solutions play a key role in enabling business and technology leaders to make the most of their data. Leading banks put data first. The winners in the Extreme Data Economy will be defined by their agility and ability to use data to power their businesses.

Resources

- [Accenture: In The Age Of Big Data, Banks Face Tricky New Challenges – Forbes](#)
- [Banking Strategy for the Long Game – Bain & Co.](#)
- [Cybercrime Costs Financial Services Sector More Than Any Other Industry – Accenture](#)
- [Exploring Next Generation Financial Services: The Big Data Revolution – Accenture](#)
- [Global Capital Markets 2018: Embracing the Digital Migration – The Boston Consulting Group](#)
- [How Big Data Analytics is Breaking Through Organizational Walls – Bain & Co.](#)
- [Rewiring Citi for the Digital Age – McKinsey](#)
- [Standard Chartered: The Force Awakens Big Data in Banking – Finextra](#)
- [2017 Compliance Risk Study: Financial Services – Accenture](#)



The journey to a next-generation advanced analytical stack begins with a single step toward the right technology. Exploring the potential impacts that an instant insight engine could have on your business doesn't require a massive investment. Start off by requesting a free demo to experience the Kinetica difference.

[Request a Demo](#)

About Kinetica

When extreme data requires companies to act with unprecedented agility, Kinetica powers business in motion. Kinetica is the instant insight engine for the Extreme Data Economy. Across healthcare, energy, telecommunications, retail, and financial services, enterprises utilizing new technologies like connected devices, wearables, mobility, robotics, and more can leverage Kinetica for machine learning, deep learning, and advanced location-based analytics that are powering new services. Kinetica's accelerated parallel computing brings thousands of GPU cores to address the unpredictability and complexity that result from extreme data. Kinetica has a rich partner ecosystem, including NVIDIA, Dell, HP, and IBM, and is privately held, backed by leading global venture capital firms Canvas Ventures, Citi Ventures, GreatPoint Ventures, and Meritech Capital Partners. For more information and trial downloads, visit kinetica.com or follow us on [LinkedIn](#) and [Twitter](#).

kinetica

Headquarters

101 California St, Suite 4560
San Francisco, CA 94111

(888) 504-7832
+1 415 604-3444

To learn more

visit us at
kinetica.com

or
[contact us](#)