Real-time Insights and Decision Making using Hybrid Streaming, In-Memory Computing Analytics and Transaction Processing

Applying Gartner's Hybrid Transaction/Analytical Processing Architecture to Cybersecurity

Welcome

Businesses are processing exponentially more data from clicks, swipes, micropayments, cyber packets, social feeds and meter readings. The volume of streaming data is overwhelming IT systems, as a result many businesses are "blind and deaf" for minutes to hours at a time. Traditional data solutions are struggling to update, organize and index their data so it can be queried in real-time to see the true state of their business.

Today's businesses need fast interactive access to tens or hundreds of terabytes of data extracted from huge data lakes and legacy systems. They also need to combine multiple data sources, both in-flight and historical data, to dynamically aggregate their data along different dimensions and visualize it in both tabular and graphical form. In particular, analysts at data driven businesses must be able to "live in their data," simultaneously interacting with both historical and real-time streaming data.

The financial services industry has been successfully dealing with similar challenges for over two decades. Banks and trading firms have coped with steadily increasing data volumes over that period using a simple scalable data architecture composed of a real-time database (RDB) and an historical database (HDB). Their approach is based on a hybrid of streaming, in-memory computing (IMC) RDBs and persistent HDBs to provide real-time services to traders, regulators and risk managers.

Analysts in other industries like energy, telecom operations systems support (OSS) and cybersecurity also need to maintain a total view of the current and past state of their grids and networks.

One of the major benefits of this architecture is that it is easy to scale up or out, simply by adding additional RDBs and HDBs. Multiple RDBs and HDBs can be provisioned for specific time periods, types of transactions, etc. In financial services it is common to find very sophisticated data architectures that combine multiple instances of streaming with complex event

processing (CEP), RDBs and HDBs with publish/subscribe capabilities, and business processing using micro-services.

The in-memory RDB addresses real-time business needs while the HDB is an immutable record of all past transactions. By querying both the RDB and the HDB one can always get a consistent view of the world and the state of the business. Increasingly, RDBs are being placed in huge non-volatile memory and HDBs are being stored on fast SSDs. Working smartly with today's memory solutions enables businesses to query all their data in near real-time providing a consistent picture directly assembled from the raw transactional data as required.

This widely used RDB/HDB architecture has lately been re-discovered by the Big Data community. Variations on the model have been described as Lambda Architecture, Gartner's Hybrid Transaction/Analytical Processing (HTAP) [1], Forrester's Translytical DB[2] and Event Sourcing. We feel the Gartner report below provides valuable guidance on how to realize business value using HTAP.

[1] Gartner, Predicts 2016: In-Memory Computing-Enabled Hybrid Transaction/Analytical Processing Supports Dramatic Digital Business Innovation by Fabrizio Biscotti, Massimo Pezzini, Nigel Rayner, Joseph Unsworth, Roxane Edjlali, Susan Tan, Errol Rasit, Andrew Norwood, Andrew Butler, W. Roy Schulte ; 17 December 2015

[2] Forrester, Emerging Technology: Translytical Databases Deliver Analytics At the Speed of Transactions by Noel Yuhanna and Mike Gualtieri; December 10, 2015

Research from Gartner:

Hybrid Transaction/Analytical Processing Will Foster Opportunities for Dramatic Business Innovation

Hybrid Transaction/Analytical Processing Will Foster Opportunities for Dramatic Business Innovation

Hybrid transaction/analytical processing will empower application leaders to innovate via greater situation awareness and improved business agility. This will entail an upheaval in the established architectures, technologies and skills driven by use of in-memory computing technologies as enablers.

Impacts

- The emergence of HTAP means IT leaders must identify the value of advanced real-time analytics, and where and how these enable process innovation.
- By eliminating analytic latency and data synchronization issues, HTAP will enable IT leaders to simplify their information management infrastructure, if they can overcome the challenges of adopting this new approach.
- Technology immaturity and established application environment value and complexity will force IT leaders to plan for long-term coexistence between HTAP and traditional approaches.

Recommendations

- Educate business leaders about HTAP and IMC concepts, and why they are important. Brainstorm with them to identify concrete opportunities to rethink business processes.
- Discuss with your strategic information management and business application providers their vision, road map and technology for HTAP implementation in their products.
- Pilot the use of HTAP architectures in individual "system of innovation" projects.
- Balance the costs of HTAP adoption (hardware and software infrastructure, migration, operational processes, and skills) with the anticipated business and IT benefits.
- Plan for coexistence and interoperability of IMC and traditional technology, mixing products from different vendors, if needed.
- Revisit your information management strategy, including governance, SLAs and life cycle management, to ensure proper organizational alignment and ownership of the information in support of using HTAP architectures along with traditional approaches.

Analysis

The notion of running transactions and analytics on the same database of record has been around since the early days of computing, but has not fully materialized so far because of a variety of issues, including technology limitations.

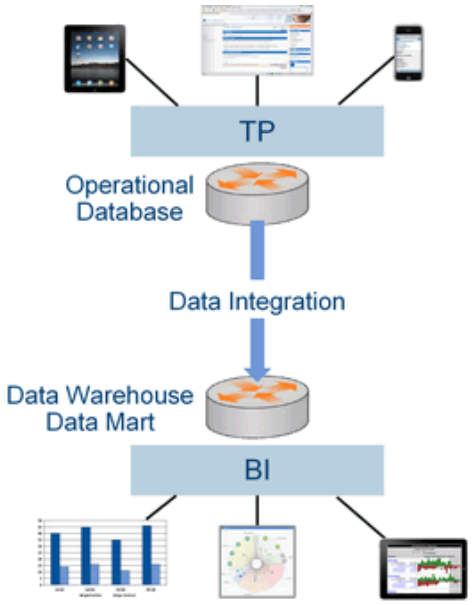
Transaction processing systems (see Note 1) are usually supported by operational DBMSs (formerly referred to as online transaction processing [OLTP] DBMSs; storing data in structures optimized for fast access to single data items ("rows").

Analytical processing systems (see Note 2) were created separately to optimize data structures for massive queries, as well as to avoid access concurrency and the associated performance and scalability impact of analytical queries over business-critical transactional systems. Such a separation also allowed the combining of data from multiple sources into analytical databases to measure enterprise performance across applications.

Therefore, transaction processing and analytical systems are usually based on distinct architectures, thus adding complexity to the information architecture and relevant infrastructure and introducing delay in data analysis (see Figure 1).

Because of technology advances such as in-memory computing (IMC), hybrid transaction/analytical processing (HTAP) architectures, which enable applications to analyze "live" data as it is created and updated by transaction-processing functions, are now realistic and possible. Although not every business process equally

FIGURE 1
Traditional Transaction and Analytical Processing



Source: Gartner (January 2014)

benefits by such architectures (for example, in certain cases latency is characteristic of the process itself and not just an induced negative effect stemming from technology limitations), HTAP will, in many cases, enable radical improvements.

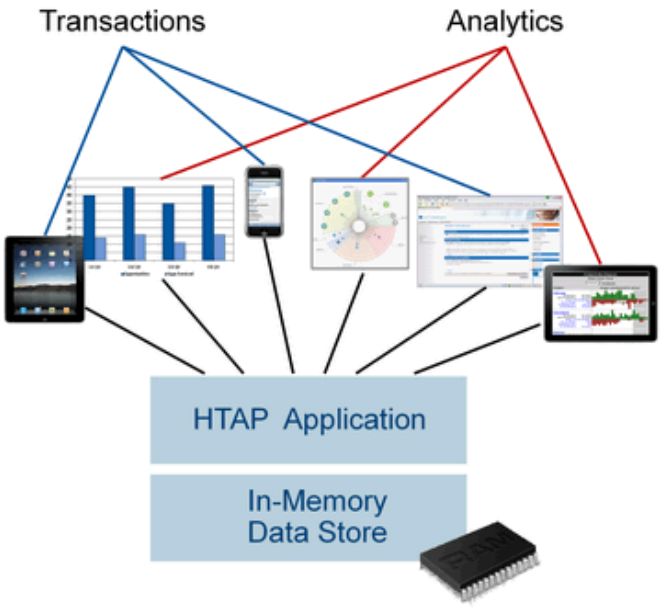
Although it is possible to build simple forms of HTAP applications using traditional DBMSs, Gartner believes that most HTAP implementations should and will be IMC-enabled. IMC technologies, such as in-memory DBMSs (IMDBMSs) and in-memory data grids (IMDGs), support a single, low-latency-access, in-memory data store that can process high volumes of transactions. These technologies can also support zero-latency analytics on that same data, including advanced analytics, such as forecasting and simulations, as well as more traditional styles of descriptive analysis (see Figure 2).

Gartner believes less than 1% of user organizations have implemented HTAP-based applications, and expects this architecture will take five to 10 years to reach mainstream adoption. Nevertheless:

- Some user organizations have already leveraged IMC-enabled HTAP to support transformative business process innovation.
- Some packaged application vendors and SaaS providers are already exploiting HTAP architectures in their offerings. Although these are in their early stages and adoption is often limited, vendor marketing efforts will soon bring this approach to business users' attention.

To support their organizations' business strategy, understanding the fundamental tenets of HTAP and defining a relevant strategy will be an imperative for IT leaders involved in the definition of application architecture, information management and business analytic strategies.

FIGURE 2
IMC-Enabled Hybrid Transaction/Analytical Processing



Source: Gartner (January 2014)

Impacts and Recommendations

The emergence of HTAP means IT leaders must identify the value of advanced real-time analytics, and where and how these enable process innovation

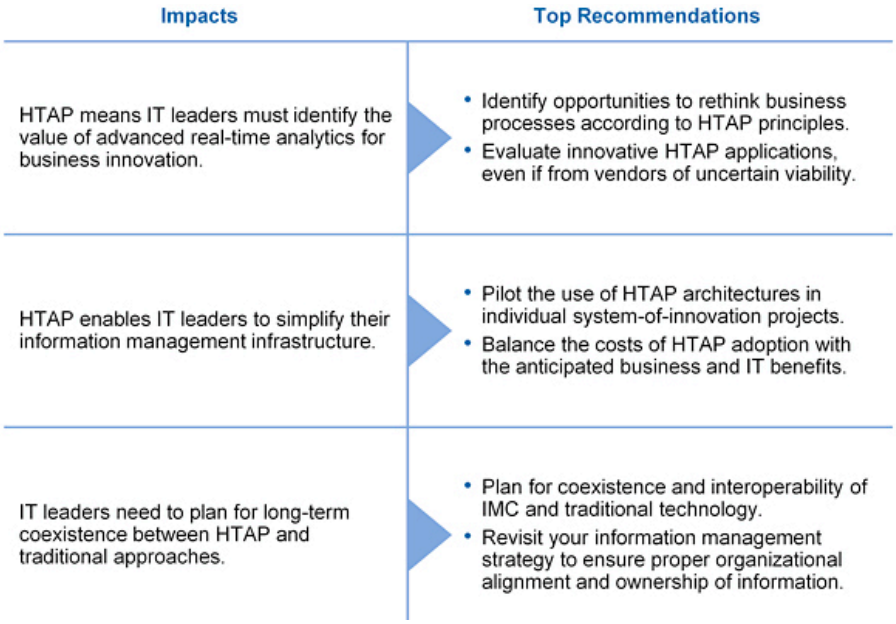
HTAP could potentially redefine the way some business processes are executed, as real-time advanced analytics (for example, planning, forecasting and what-if analysis) becomes an integral part of the process itself, rather than a separate activity performed after the fact. This would enable new forms of real-time business-driven decision-making processes (see Note 3 for some examples). Ultimately, HTAP will become a key enabling architecture for intelligent business operations.

HTAP will enable business leaders to perform, in the context of operational processes, much more advanced and sophisticated real-time analysis of their business data than with traditional architectures. Large volumes of complex business data can be analyzed in real time using intuitive data exploration and analysis without the latency of offloading the data to a data mart or data warehouse. This will allow business users to make more informed operational and tactical decisions.

Running more complex diagnostic and predictive analytics in real time (and leveraging very granular transactional data) on the database of record will allow decision makers to be alerted to business trends and situations that may require their immediate attention. HTAP architectures accomplish this by improving business leaders' situation awareness in operations (for example, for applications such as risk management and fraud detection), and by providing constantly updated forecasts and simulations of future business outcomes.

FIGURE 3

Impacts and Top Recommendations for HTAP for Business Innovation



Source: Gartner (January 2014)

However, figuring out how HTAP can help process innovation is challenging. Business leaders need to think outside of the box to determine how this architecture could transform processes rather than just provide existing styles of analytics faster and without latency. But application architects and developers (and also system integrators and packaged/SaaS application vendors) have little or no experience with how to design, implement, deploy and operate HTAP applications. Hence, it may be difficult for them to assess whether the innovation pushed by business leaders is really possible to implement.

However, IT leaders don't necessarily need to adopt a big-bang approach to HTAP. Use of IMC technologies provides opportunities for user organizations to adopt HTAP architectures incrementally, through a trial-and-error approach. For example, IM analytics have enabled data discovery capabilities, allowing business users to freely (that is, without much support from the IT department) interact with data. With these tools, business users can explore data and discover new insights without any help from IT. The data used is often the result of a data mashup created by business users outside any governance process. As a result, such a mashup adds yet one more copy of the transactional applications' data. With HTAP, transactional data is directly available for discovery and creating copies is not needed.

This incremental approach to HTAP favors adoption of a rapid and agile method to new application function delivery. It is not necessary to develop detailed requirements or to implement all the analytical features before putting the new system in production. In many cases, it is possible to initially implement the core set of capabilities, incrementally add new analytics, and then let end users develop or customize (by themselves) the features they deem necessary.

Ultimately, HTAP architectures will do for business analytics what data visualization tools have done for query and analysis: put the right tools in the hands of the business users to perform their own analytics, without needing IT support and without creating separate, stand-alone "spreadmarts" to perform those analytics that IT is not providing.

Recommendations:

- Educate business users about the concepts of HTAP and IMC, and why these are important new approaches using simple metaphors (see Note 4).
- Brainstorm with business leaders to identify opportunities to rethink business processes, from an HTAP perspective, to support system-of-innovation initiatives via greater situation awareness, advanced real-time analytics and business-goal-driven decision making.
- Evaluate native HTAP-packaged/SaaS applications targeting process innovation, even those from small vendors of uncertain viability. Don't wait for the megavendors to deliver if you aim for a competitive advantage.

By eliminating analytic latency and data synchronization issues, HTAP will enable IT leaders to simplify their information management infrastructure, if they can overcome the challenges of adopting this new approach

HTAP will help simplify organizations' information management infrastructure by eliminating, or at least reducing, the need to duplicate data and to keep data consistent. HTAP architectures also imply a redefinition of the role of traditional data warehouse architectures.

However, the degree of simplification enabled by HTAP will vary. By not requiring a data mart or data warehouse, HTAP could make it simpler to meet reporting and analytical needs based on operational data from a single instance of an application

stack (such as an ERP suite). A small number of organizations are already realizing these benefits. However, HTAP will not have such a beneficial impact on large and complex organizations that need to aggregate data from multiple sources (for example, SaaS applications, cloud information services, social networks and on-premises applications). However, even in these organizations, data marts created just to support analytical and reporting needs from a single application could be eliminated, which will simplify some of the information management landscape.

By definition, at least in the context of a single application stack, HTAP architectures address the four major drawbacks of traditional approaches:

- **Architectural and technical complexity.** In traditional approaches, data must be extracted from the operational database, transformed and loaded into the analytical database, which requires the adoption of database replication; extraction, transformation and loading (ETL) tools; enterprise service buses (ESBs); message-oriented middleware (MOM); and other integration tools.

  In HTAP, data doesn't need to move from operational databases to separated data warehouses/data marts to support analytics.

- **Analytic latency.** In a classic setting, it can take hours, days or even weeks from the moment data is generated by the transaction processing application to when it can be used for analytics. Although this is adequate for certain types of analytics, and even processes, it may be suboptimal for others. For example, being able to perform financial consolidation at any point in the month can enable a CFO to better evaluate the business impact of economic trends and take early corrective actions.

  The transactional data of HTAP applications is readily available for analytics when created.

- **Synchronization.** If analytical and transactional data storage is separated, when business users want to "drill down" from a point-in-time aggregate into the details of the source data in the operational database, in many cases they find the source of data "out of synch" because of the analytic latency.

  In HTAP, drill-down from analytic aggregates always points to the "fresh" HTAP application data.

- **Data duplication.** In a traditional architecture, multiple copies of the same data must be administered, monitored, managed and kept consistent, which may lead to inaccuracies, timing differences and inconsistency.

  In HTAP, the need to create multiple copies of the same data is eliminated (or at least reduced).

However, adopting HTAP has its challenges: The concept is immature, industry experience is still limited to the most-leading-edge organizations in a few industry sectors (primarily financial services), best practices are not yet crystallized, the vendors' landscape is still quite turbulent, and relevant skills are almost impossible to find.

IT leaders also will face the issues associated with the use of IMC technology, which is a key enabler for HTAP initiatives:

- Often, the code of traditional applications must be re-engineered to take maximum advantage of IMC technologies and to enable integration with advanced analytics tools.

- Migrating to the IMC-enabled version of a traditional packaged application doesn't necessarily lead to an HTAP architecture if the vendor didn't go through the process of optimizing its product code for IMC and adding real-time analytics.

- In-memory data management technologies (IMDBMS and IMDG) are not fully proven in support of HTAP use cases.

- Data governance policies and processes will have to be revised to cover in-memory data. For example, data quality checks need to happen in the application itself, rather than when data is being moved into the data warehouse.

- A combination of data sources in the HTAP in-memory data store may be required to perform advanced analytics. This is a challenge in its own right, but it is further exacerbated by the fact that data integration tools that feed in-memory data stores from other data sources are still few and are not fully proven in real-life production deployments.

Therefore, HTAP adoption may require significant costs and risks, which would make it hard to justify only on the basis of information and application infrastructure cost reductions, although these considerations should be factored into the business case for HTAP-based application projects.

Recommendations:

- Discuss with your strategic data management and business application providers their vision, road map and technology for HTAP implementation in their products.

- Experiment with IMC technologies to assess the suitability of these products for your HTAP requirements.

- Pilot the use of HTAP architectures in individual, high-risk/high-reward system-of-innovation projects (as opposed to enforcing HTAP as an architectural principle across the board). This will help identify challenges that will need to be addressed

once the high-impact potential use cases have been identified as candidates for a more extensive deployment.

- Evaluate the costs of HTAP adoption (skills acquisition, application and hardware infrastructure refresh, application migration, and operations and management process updates), and balance those with the anticipated business and IT benefits.

Technology immaturity and established application environment value and complexity will force IT leaders to plan for long-term coexistence between HTAP and traditional approaches

Gartner believes HTAP adoption will grow significantly over the next five years because of its significant business impact. Nonetheless, IT leaders should plan for a coexistence of traditional and HTAP-based systems. As discussed previously in this research, data warehouse architectures will remain necessary to support extended analysis that involves large amounts of historical data/big data that comes from multiple internal and external, structured and unstructured data sources. Individual HTAP systems will be contributors to those (logical or physical) data warehouses, but will not fully replace them.

Moreover:

- In many instances, a traditional approach meets business requirements, and migrating to an HTAP architecture would not be justified and, often, not even desirable or possible.
- Even if theoretically possible, it will be practically impossible to migrate many established applications toward HTAP due to the massive organizational and technical efforts, and because of the inertia of packaged business application vendors.

Recommendations:

- Plan for coexistence and interoperability of traditional and HTAP-enabling technologies, possibly procuring products from different vendors, if needed. A single-supplier strategy is premature, as today, none can provide a comprehensive and mature portfolio of HTAP-enabling products.
- Identify the role of HTAP in the evolution of your data warehouse toward the logical data warehouse.
- Revisit your information management strategy, including governance, SLA and life cycle management, to ensure proper organizational alignment and ownership of the information in support of HTAP architectures along with traditional approaches.

Note 1

Transaction Processing Application Systems

These systems:

- Enable fast access to data to support business processes, such as order entry, banking operations, travel reservation systems, e-commerce and myriad other day-to-day activities in every industry sector.
- Collect, process and manage data in support of critical business processes.
- Must support the SLAs that are defined by the business and ensure that all transactions are processed reliably.

Note 2

Analytical Processing Systems

These systems:

- Support efficient analysis of data for reporting, business intelligence and other initiatives that require the fast scanning of large databases to create data summaries and aggregations.
- Address organizations' needs to monitor the business, measure performance and identify trends by combining multiple sources of data from multiple applications.

Note 3

Examples of How HTAP Can Improve Business Processes

HTAP can improve business processes by:

- Deciding on how to fulfill a purchase order on the basis of its impact on profitability or customer satisfaction.
- Using predictive analytics to inform decision makers that business and financial targets are unlikely to be met by financial period end, so they can take corrective action (instead of providing diagnostic analysis after the event).
- Notifying logistics companies that shipments may miss their delivery timelines. Thus, business analysts can adjust the itineraries or renegotiate shipping alternatives with clients in real time via what-if analysis.

Note 4

Example of How to Sell HTAP to the Business

One CIO was able to "sell" HTAP to his business executives by suggesting that, while their traditional BI approach allowed them to instantly see what happened yesterday, HTAP would give them real-time visibility into what was happening now.

Source: Gartner Research Note, G00259033, Massimo Pezzini, Donald Feinberg, Nigel Rayner, Roxane Edjlali, 28 April 2015

Proven Performance and Flexibility – Kx HTAP Solution Cybersecurity Use Case

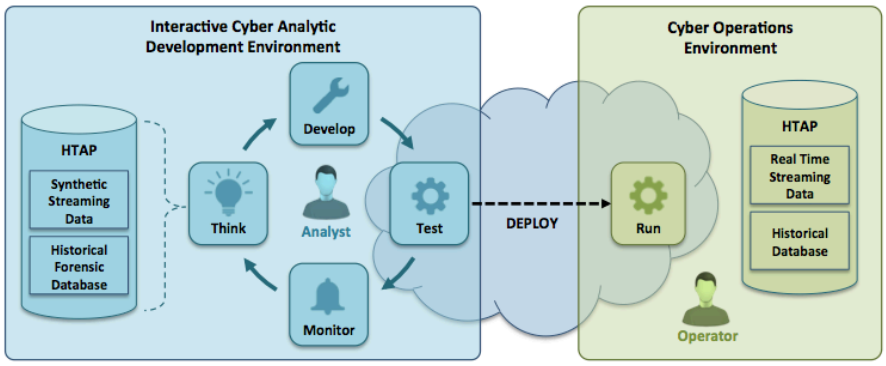
Kx's implementation of HTAP architecture provides businesses with the means to make informed business decisions based on analytics run over real-time, streaming and historical data. As Gartner says, "HTAP will enable business leaders to perform, in the context of operational processes, much more advanced and sophisticated real-time analysis of their business data than with traditional architectures." [1]

One of the challenges in deploying IMC solutions is the integration of streaming, in-memory and storage based processing which often requires the use of different vendors and technologies. Another challenge is the need to deal with burst of traffic associated with black Friday shopping or a cyber attack which may exceed the capacity of the in-memory DB or grid. Kx addresses the first by providing a single technology which handles streaming, in-memory and disks based databases greatly simplifying HTAP implementation. Unlike many in-memory only solutions Kx technology has been in use for decades to handle both in-memory and disk-based workloads hence can accommodate burst requirements.

In order to meet the demanding needs of capital markets Kx is designed for performance on modern processors with the compact database engine staying in the instruction cache, and column processing leveraging memory prefetch and vector instructions.

This means that Kx can handle more data at once and is more efficient at aggregating and analyzing data. Proven in one of the most demanding application areas, financial services, industries struggling to address increasingly complex regulatory and business decision requirements are turning to Kx's HTAP solution.

FIGURE 1
Sample Kx - HTAP Development Environment



Source: KX

Source: Kx

[1] Gartner, Hybrid Transaction/Analytical Processing Will Foster Opportunities for Dramatic Business Innovation, Massimo Pezzini | Donald Feinberg | Nigel Rayner | Roxane Edjlali, 28 April 2015

Analyst for Kx

The Kx for Surveillance solution provides HTAP surveillance capabilities to the financial markets where determination of rogue activity is based as much on historical data (eg: transactions, news events, order patterns) as the immediate transaction under investigation. In addition, Analyst for Kx is a robust solution that implements Gartner's HTAP architecture for cyber-analytics and telecommunications and has been providing timely information for enforcement activities for several years.

There are few areas where the competition is so intense, and the stakes are as high as cybersecurity. Cybersecurity is a high stakes game that is always changing, where humans augmented by computers continually seek new techniques to both obfuscate their attacks and mask their presence in target systems. "Both defenders and attackers collaborate with their respective peers to gain competitive advantage. Both review the past to anticipate the future. Both constantly scan software to identify potential defects maintaining an inventory for the future repairs/exploits."

Classic cyber-defense strategies for Individuals and corporations count on detecting known attacks and using well-known digital fingerprints to block potential threats. Experts will confidently assert that no one is immune from attacks, and daily news stories bear them out. Assume that you will be hacked, and plan now to focus on identifying and countering attacks.

When a cyber-attacker gains access they often hide inside target systems for many months, gathering information about the system's software and hardware in order to eventually mount a more effective attack. In fact, most sophisticated attacks are built up over long periods of time, hiding innocently in an otherwise normal stream of internet traffic, remaining undetected by classic cyber-approaches that are based on simple signatures or looking at traffic in a small sliding window. These old-style defenses are insufficient in today's world. In order to detect modern attacks one needs to look simultaneously at both current as well as all past traffic. Forensic analysis on past data will often expose full or partial patterns that often appear again in future traffic.

Analyst for Kx is a product designed to help human cyber-analysts anticipate and detect new forms of attack. Part of a suite of surveillance products, Analyst for Kx provides the tools necessary to deal with ingesting, transforming and visualizing internet data. It is one of the only tools that enable experts to interactively explore huge volumes of both current and past traffic. Only by thorough forensic analysis of historical data is one able to understand attack patterns and characterize normal behavior, both being essential to building and validating usage and traffic models. Kx also has significant experience with stock market surveillance, which is a very similar cyber-crime and uses similar technical approaches.

In order to cope with future threats, cyber-analysts place themselves in the mindset of an attacker, creating models of potential attacks and devising associated detectors for such attacks. These point of decision models run in the live stream of traffic setting off an alarm when a detector identifies a potential attack. Analysts then quickly verify if there is a cause for concern and if so, immediately take steps to counter the attack.

As a practical matter designing and implementing model-based behavioral cyber-security requires a full programming language and associated development tools, together with comprehensive analysis and visualization tools. In addition, machine learning and statistical libraries that support a full range of statistical and deterministic modeling are essential. Efficient probabilistic reasoning techniques to deal with the realities of fuzzy/uncertain reasoning about noisy data are also needed. Analyst for Kx offers all of these capabilities, seamlessly integrated with the proven Kx HTAP solution.

Source: Kx