**Memory-Driven Computing designed for Next-Gen In-Memory Analytics Solutions**

**Current enterprise challenges with handling growing volume of data**

Businesses are experiencing an exponential increase in data, coming from an explosion of sources and we have a vanishingly small time to turn that data into meaningful action. In past, the fact that data is double every two years led to a very famous law of electronics i.e. Moore’s law which stated that transistors will double in integrated circuit every 2 years. This prediction is already proving wrong as data growth and data analytics requirements are outpacing the compute and storage technologies that provided the foundation of processor-driven architectures for the last five decades.

Deriving such time critical insights, requires architectural shift from compute-driven analytics to memory-driven analytics. This architectural redesign is driven by following challenges and experiments in recent times.

Memory-Driven Computing is an almost infinitely flexible and scalable architecture that can complete any computing task much faster, using much less energy, than conventional systems. The performance of Memory-Driven Computing is possible because now any combination of computing elements can be composed and communicate at the fastest possible speed – the speed of memory.

Only through a new architecture like Memory-Driven Computing will it be possible to simultaneously work with every digital health record of every person on earth, every trip of Google’s autonomous vehicles and every data set from space exploration all at the same time—getting to answers and uncovering new opportunities at unprecedented speeds.

**In this blog series**, I am going to highlight key New-Gen Analytics workloads taking advantage of Memory-Driven Computing architecture and scaling for real-time analytics requirements.

**Next-Gen In-Memory Analytics Solution leveraging Memory-Driven Computing Infrastructure**

Let us look at few industry use-cases that take advantage of architectural principals of compute- driven analytics.

**Use-Case-1: Cybersecurity Analytics:** Cybersecurity Analytics involves identifying cyber intrusion behaviours in a deployed infrastructure comprising of complex network of servers, routers, gateways, storage etc. Responding quickly to potential threats requires security tools capable of analysing billions of threat signals in real-time.

Developing such Cybersecurity Analytics involves analysing massive volume of Infrastructure network traffic information from network connection logs, http logs, dhcp logs, smtp logs, netflow information etc. and there by establish a network of infrastructure entities and relationships. This is achieved by building a network graph and following which network anomaly patterns are recognized leading to identifying threats like Zombie reboot, RDP hacking etc.

Typical size of these network graphs comprises of **billions of graph nodes and properties and relationship between graph nodes**. Deriving anomaly patterns across these billions of nodes in real-time requires existence of entire graph in-memory with TBs of large memory infrastructure.

Cybersecurity Analytics use-case with Graph Databases is very ideal for memory-driven analytics.

**Use-Case-2: Real-Time Recommendation Engine:** In this day and age, the need to build scalable Real-Time Recommendations is increasing day by day. With more internet users using e-commerce sites for their purchases, these e-commerce sites have realized the potential of understanding the patterns of the users' purchase behaviour to improve their business, and to serve their customers on a very personalized level. To build a system which caters to a huge user base and generates recommendations in real time, we need a modern, fast scalable system.

Deploying a Real-Time Recommendation engine involves In-Memory data processing unifying user data from social-media data sources, existing customer management solutions, existing warehouse historical data etc. Building such unified analytics platform can be achieved with a combination of Spark based In-Memory data processing and transformed representation of customer data in graph models. These unified analytics built with Spark involves application of massive transformation and action. These phases of transformation and action undergo massive data shuffle operation to pre-process which is very costly in cluster operation leading to lack of real-time capabilities.

Typical scale of shuffle spans across TBs of data undergoing Repartition, GroupBy and Join operations which are very expensive owing to frequent I/O operations. Making these TBs of data available In-Memory with large memory infrastructure enables Spark to perform data pre-processing operations faster to meet the requirements of **Fast Data Analytics.**

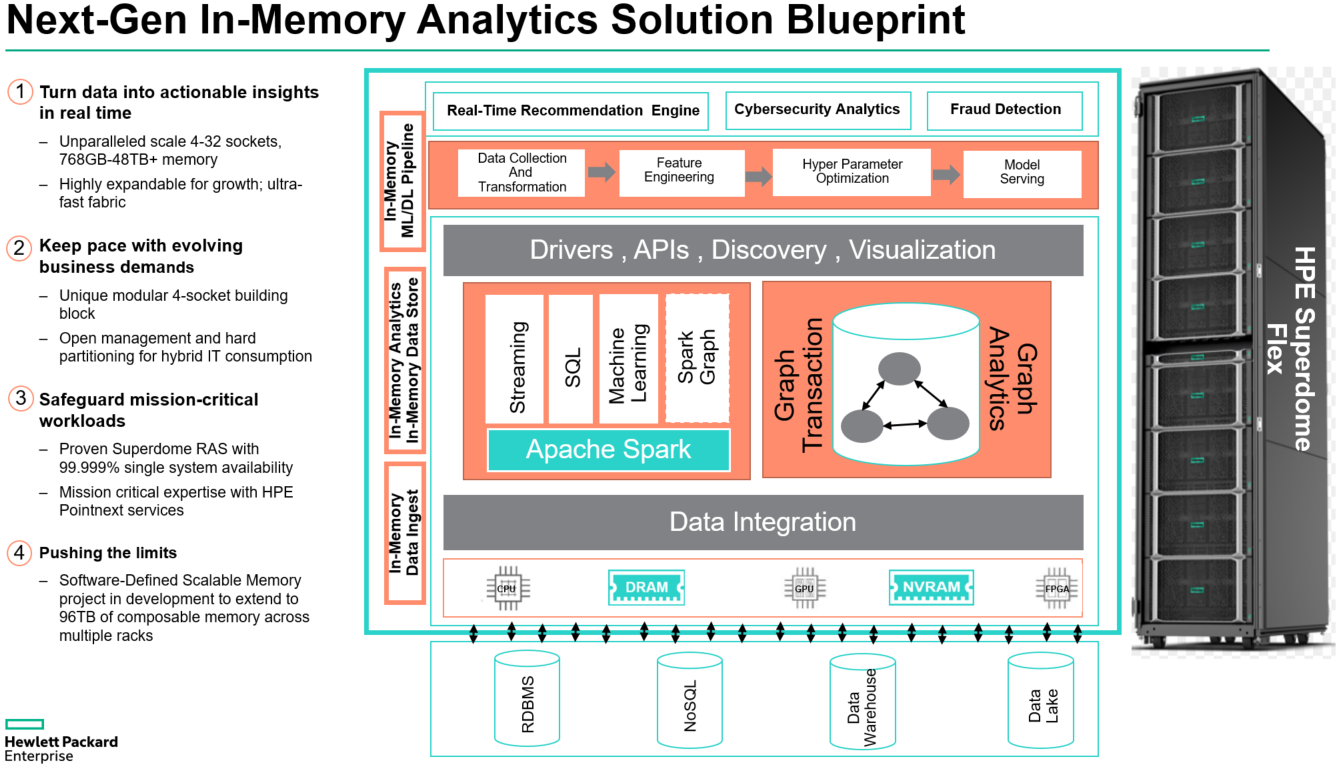
Real-Time Recommendation Engine with Fast Data Analytics is yet another ideal use-case for memory-driven analytics.

**Use-Case-3: Fraud Detection:** Fraud management has been known to be a very painful problem for banking and finance firms. Card-related frauds have proven to be especially difficult for firms to combat. Technologies such as chip and PIN are available and are already used by most credit card system vendors, such as Visa and MasterCard. However, the available technology is unable to curtail 100% of credit card fraud.

Building a fraud prevention solution requires analysing credit card transaction in sub-millisecond time frame, detecting outliers in which data-set is verified to identify potential anomalies in the data. With the rise of **machine learning** (**ML**), **artificial intelligence** (**AI**), and **deep learning**, it becomes feasible to analyse massive volume of transactions feeding into enterprise credit card network. These machine learning models are first trained against historical transactions and live inference is achieved by building machine learning pipeline for data acquisition, feature engineering and model serving.

Achieving Real-Time fraud detection against streaming credit card transactions requires building machine learning pipeline in-memory to perform data collection and transformation, feature engineering, hyper-parameter optimization and model serving to make high precision predictions. This can be achieved by implementing Spark and Tensorflow with TensorFrames and taking advantage of co-located GPUs for faster model training.

Building upon these use-cases, below is the high-level architecture for Next-Gen In-Memory Analytics Solution designed to take advantage of scale-up Memory-Driven Computing Infrastructure with HPE Superdome Flex.



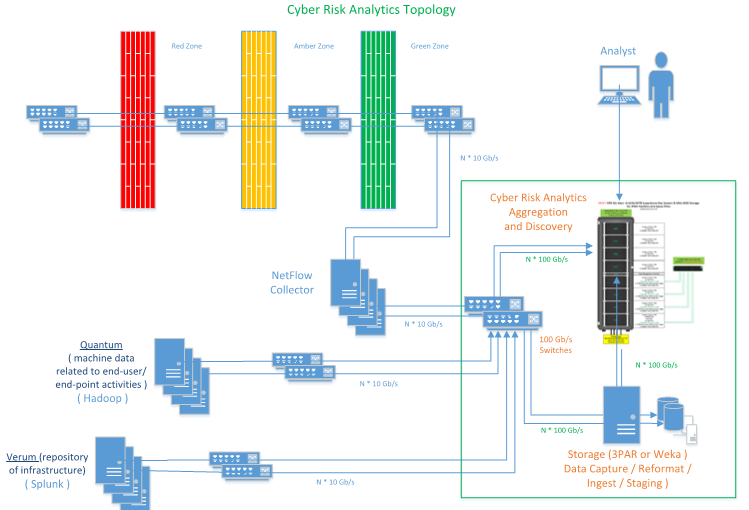
**Key features of In-Memory Analytics solution are**

1. Adopt and build **large scale In-Memory data pipeline** with Spark taking advantage of large memory off-heap cache for shuffle operations for data preparation
2. **Accelerate machine learning training and inference** against massive volume of train and test data and achieve high precision accuracy for predictive analytics
3. **Accelerate connected data analytics leveraging Graph Databases** using search algorithms like Page Rank or Single Source Shortest Path (SSSP) taking advantage of large memory Superdome Flex infrastructure.

Let us look at how In-Memory analytics is achieved leveraging Memory-Driven Computing for real customer implementations.

A large Financial Services organization implemented Cybersecurity Analytics to identify cyber threats in their operational network by detecting network hacks, zombie reboots, RDP hacking etc.

In order to establish this Cybersecurity Analytics, network related data was captured over 90 days period. These data-sets included network connection logs, http logs, dhcp logs, smtp logs, netflow information, packet capture (pcap) data from the operational network. Historical infrastructure operational data was also available in existing Hadoop based Data Lake.



First challenge was to integrate all this data. This was achieved by aggregating all the data from these multiple sources into large memory infrastructure hosting terabytes of data.

Subsequently, aggregated data was transformed into a Graph Data Model and a network graph was built to represent these network entities. Size of network graph included 20 Billion Graph edges (17.9 Billion Netflow Edges and 1.5 Billion log edges) and 212 billion graph edge properties against 3TB of input data from network.

In order to build such network graph, an enterprise software product was implemented in HPE Superdome Flex large memory environment to read the input data from multiple data sources as highlighted above and the tool generated a powerful network graph with billions of edges and edge properties.

Finally, pattern matching operation was performed to detect anomalies. This was achieved by developing complex pattern matching queries. Query response time was measured with deployed in-memory graph and compared against scale-out cluster based deployment and results indicated 1000x performance improvement.



Thus, HPE Superdome Flex Memory-Driven Computing infrastructure enables Next-Gen In-Memory Analytics solution for critical use-cases like Real-Time Recommendation Engines, Security Analytics and Fraud Detection across multiple industries.

**References**

*[1] 6-cybersecurity Mega Trends*

*<https://www.hpe.com/us/en/insights/articles/6-security-megatrends-1905.html>*

*[2] Scale-Up Graph Analytics with Reservoir Labs and Trovares:*

*<https://www.reservoir.com/wp-content/uploads/2019/06/ENSIGN-SDFlex-WP.pdf>*

*[3] Memory Drive Computing in Superdome Flex:*

*<https://www.hpe.com/us/en/newsroom/blog-post/2017/05/memory-driven-computing-explained.html>*

*[4] Mission critical Infrastructure for Data Driven Enterprise*

*<https://www.hpe.com/hpe-external-resources/a00037000-7999/enw/a00037029?resourceTitle=Mission-critical+infrastructure+for+the+data-driven+enterprise>*

**Ashish Kumar**

**Data and Analytics Solutions Architect**



**David J. Furtzaig**

**Director, Chief Technologist**

