GridGain - Solving data immediacy problems for real-time decisions

The average enterprise stores around 10 PB of data made up of files, images, audio, video, text, structured data and unstructured data. Of that, only 15% is considered to be business-critical data. However, that data may be spread across multiple data centers, cloud repositories, edge data stores, or remote locations. Even within a data center, such critical data may be stored across multiple data warehouses, databases, file systems, or other such storage mechanisms. This makes real-time business decision-making very difficult as all the key data points may not be available for immediate use. Enterprises struggle with the need for data immediacy.

Data Immediacy

Data immediacy is the requirement to get real-time access to all the key data points to make mission-critical business decisions that truly impact business performance, customer retention or other such key metrics. This is not just about latency pertaining to data access. While that is relevant, there are other aspects such as context, correlation, data store location, data quality etc., that will impact the decision-making process.

For example, when trying to detect fraud in banking transactions, instantaneous decision-making is critical to prevent significant monetary damages to the bank or the account holder. While your decision engine may have instant access to the incoming banking transactions in real-
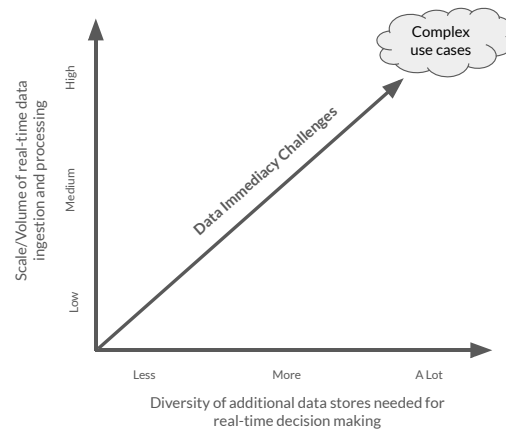
time, the real challenge may be getting immediate access to the specific account's recent transaction history, location of the account holder, recent travel purchases, previous fraud alerts for this account etc. All of these data points may be spread across multiple data stores and getting all of that in the context of the current transaction and applying correlation can take time. This delay can cost the business dearly by not being able to detect the fraud in real-time.

Challenges with current data platforms

A lot of the real-time data platforms in the market offer the key capabilities of data ingestion, real-time processing, data movement, and eventually analytics. They do not solve the problem of the data silos that get created after the data is ingested into the enterprise. The data store vendors, on the other hand, solve for bringing all relevant data into one store but fall short of addressing data immediacy issues, as explained above. Enterprises are looking for a platform that not only offers the ability to ingest real-time data but also enables getting real-time access and compute to other data stored across the enterprise to make instant business decisions.

GridGain

GridGain offers a unified data platform that is designed to deliver high-performance, real-time data processing and analytics at scale. Based on a distributed memory-first architecture, it can horizontally scale across multiple nodes, allowing it to process large volumes of data and high transaction loads. Combining that with its colocated compute capabilities, GridGain delivers real-time data processing and analytics at low-millisecond latencies.

Diagram: Scale/Volume of real-time data ingestion and processing (Low, Medium, High) vs Diversity of additional data stores needed for real-time decision making (Less, More, A Lot). Arrow labeled "Data Immediacy Challenges" pointing toward "Complex use cases".

GridGain sets itself apart from the competition by solving for data immediacy challenges and delivering on some of the key capabilities such as -

- **In-Memory computing:** From its roots as an in-memory data grid, GridGain plays to its strengths by not only allowing for storage access to memory, disk, and cloud stores but also offering colocated compute that enables you to run compute code and application logic in the same memory space where your data is.

- **Heterogenous data integration:** GridGain supports easy integrations into a wide range of data sources like RDBMS, NoSQL databases, data lakes, data warehouses and even streaming sources such as Apache Kafka. This enables a single view of all that data, enabling data immediacy.

- **Distributed Architecture:** GridGain's inherently distributed architecture offers massive horizontal scalability on-premises or in the cloud, supporting high availability and strong data consistency.

- **Cross-platform deployability:** GridGain runs on practically any cloud platform, on-premises, as a managed service, in containers, and even on Mainframes

- **Diverse API support:** GridGain allows for data access via ANSI SQL, continuous queries or key-value APIs. In addition, it also supports programmatic access with support for APIs across Java,C#/.NET, C++, Python, and REST.

- **Comprehensive management and monitoring:** GridGain Control Center offers a complete solution for troubleshooting, monitoring, visualization, and performance optimization of your GridGain clusters.

Key use cases for GridGain

- **Real-time Risk Management:** Financial institutions leverage GridGain to analyze streaming data (transactions, market trends) from multiple sources for real-time financial portfolio risk assessment, determining optimal hedging strategies, or transactional fraud detection. With in-memory processing, GridGain enables split-second decisions to mitigate risk and helps identify and prevent fraudulent incidents before they occur.

- **CDR Management and 5G:** Telcos can use GridGain to manage high volumes of Call Data Records (CDRs), making the data available to all their business and operational systems. With the advent of 5G use cases, speed, performance, and extremely low latency are much needed in a data platform. GridGain is ideally suited for that.

- **Real-time Customer 360° View:** Retailers can gain a unified customer view in real-time using GridGain. It combines data from various sources (engagement patterns, purchase history, loyalty programs) to personalize offers and recommendations, enhancing customer experience.

- **Digital Integration Hub:** GridGain acts as a central hub, seamlessly integrating data from various sources (databases, data lakes, and even the mainframe) and making it accessible to applications. This eliminates data silos and fosters a unified data environment.

GridGain offers a unified data platform for real-time data processing, storage, and analytics to solve data immediacy challenges across various industry verticals. By including heterogeneous data integration options and a diverse set of APIs, GridGain offers a lot of flexibility in creating a data hub to combine a variety of data sources across the enterprise in real-time. GridGain is also purpose-built for security, scalability and performance to address the most demanding data processing use cases.

Dinesh Chandrasekhar has been a data management veteran, thought leader, and data practitioner for 30+ years. As the chief analyst and founder of Stratola, he speaks and writes on the latest topics in data-in-motion, real-time analytics, IoT, Observability, GenAI, and more. Follow his work at www.stratola.com.