GridGain: In-Memory Computing for the Internet of Things

Over the last decade, Internet of Things (IoT) adoption has exploded. Today there are more connected devices than people. By 2020, Gartner expects the IoT to have over 20 billion connected things, a conservative estimate compared to other analysts. The information generated by connected devices requires an enormous amount of real-time processing and storage. Some connected cars now upload 25GB of data per hour. A connected airplane can generate a petabyte a flight. But to realize the benefits of IoT, you need to choose the right architecture and set of technologies that can process large data streams, identify important events and react in real-time.

Using GridGain for Connected Cities

A leading IoT company for connected utilities, buildings and cities with hundreds of millions of connected devices was struggling with the scalability and costs of its existing infrastructure. Its inability to scale was preventing it from supporting larger provider and municipality customers.

By using GridGain as the core of their IoT infrastructure to ingest data from their existing IoT networks, the company was able to leverage GridGain's in-memory speed and unlimited linear horizontal scale to support the largest customers. GridGain also helped the company reduce overall operating expenditures by over 20%.

Many companies have succeeded with IoT, and solved their challenges around speed, scalability and real-time analytics using GridGain® and Apache® Ignite™. GridGain is the leading in-memory computing platform for real-time business. It is built on Apache Ignite, one of the top 5 Apache Software Foundation (ASF) projects for the last two years. GridGain includes enterprise-grade security, deployment, management and monitoring features which are not in Ignite. GridGain Systems contributed the code that became Ignite to the Apache Software Foundation and continues to be the project's lead contributor.

GRIDGAIN PLATFORM FOR IOT

An increasingly common architecture for IoT data processing and applications is a combination of Apache Kafka® for messaging from the IoT network; GridGain or Apache Ignite for data ingestion and management; and Apache Spark™ for stream processing. GridGain ingests Kafka data streams, stores and manages the data, prepares the information for Spark, and performs real-time analytics and decision automation to react to events as they occur. With GridGain all stream and batch processing for transactions, analytics and any decision automation can be performed through a single in-memory computing layer in real-time. Gartner and others call this approach hybrid transactional/analytical processing (HTAP).

GridGain and Ignite are used by the largest companies in the world to ingest, process, store and publish IoT and other streaming data for large-scale, mission critical business applications. They are used by several of the largest banks in the world for trade processing, settlement and compliance; by telecommunications companies to deliver call services over telephone networks and the Internet; by retailers and e-commerce vendors to deliver an improved real-time experience; and by leading cloud infrastructure and SaaS vendors as the in-memory computing foundation of their offerings. Companies have been able to ingest and process streams with millions of events per second on a moderately-sized cluster.

STREAM INGESTION AND DATA PREPARATION

GridGain can ingest, store, process, analyze and publish large volumes of streaming data with low latency, unlimited scalability and high availability. It has been used to ingest millions of events per second on a moderately-sized cluster. GridGain is widely used with major streaming technologies including Apache Camel, Apache Flink™, Apache Flume™, Apache Kafka, Apache RocketMQ™, Apache Spark, Apache Storm, Java Message Service (JMS), MQTT, Twitter® and ZeroMQ to ingest, process and publish streaming data. Once loaded into the cluster, developers can leverage built-in GridGain libraries for concurrent data processing, including distributed SQL queries and machine and deep learning. Clients can also subscribe to continuous queries which execute and identify important events as streams are processed.

SPARK SUPPORT

Ignite also provides the broadest in-memory computing integration with Apache Spark. The integration includes native support for Spark DataFrames, an Ignite RDD API for reading in and writing data to Ignite as mutable Spark RDDs, optimized SQL, and an in-memory implementation of HDFS with the Ignite File System (IGFS). When deployed together, Spark can access all of the in-memory data in Ignite, not just data streams; share data and state across all Spark jobs; and take advantage of all of GridGain's in-memory loading and processing capabilities including its distributed SQL, advanced indexing and other massively parallel processing (MPP) capabilities. GridGain can improve Spark SQL query performance by up to 1000x.

MACHINE AND DEEP LEARNING

GridGain's Continuous Learning Framework enables companies to add machine and deep learning for intelligent decision automation. With GridGain, companies can train models against massive data sets and re-run training in "mid-stream" during stream processing to improve the models based on the latest data. GridGain provides several standard machine learning algorithms optimized for MPP including linear and multi-linear regression, k-means clustering, decision trees, k-NN classification and regression. It also includes a genetic algorithm and multilayer perceptron for deep learning. Developers can develop and deploy their own algorithms across any cluster as well using the Compute Grid.
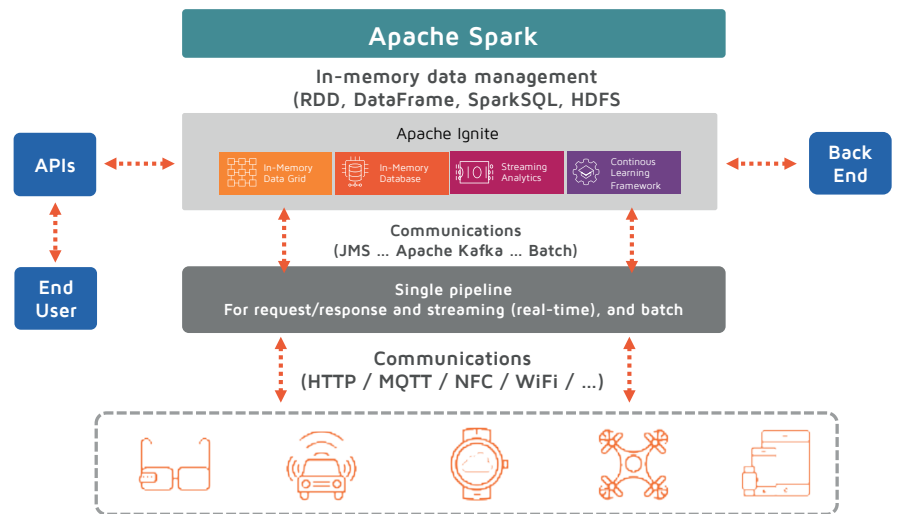
Learn More

Introducing the GridGain In-Memory Computing Platform

Contact GridGain Systems

To learn more about how GridGain can help your business, please email our sales team at sales@gridgain.com, call us at +1 (650) 241-2281 (US) or +44 (0)208 610 0666 (Europe), or complete our contact form at www.gridgain.com/contact and we will contact you.

About GridGain Systems

GridGain Systems is revolutionizing real-time data access and processing by offering an in-memory computing platform built on Apache Ignite. GridGain solutions are used by global enterprises in financial, software, e-commerce, retail, online business services, healthcare, telecom and other major sectors, with a client list that includes ING, Sberbank, Finastra, IHS Markit, Workday, and Huawei. GridGain delivers unprecedented speed and massive scalability to both legacy and greenfield applications. Deployed on a distributed cluster of commodity servers, GridGain software can reside between the application and data layers (RDBMS, NoSQL and Apache® Hadoop®), requiring no rip-and-replace of the existing databases, or it can be deployed as an in-memory transactional SQL database. GridGain is the most comprehensive in-memory computing platform for high-volume ACID transactions, real-time analytics, web-scale applications, continuous learning and HTAP. For more information, visit www.gridgain.com.