

A night-time cityscape with numerous skyscrapers and buildings, all illuminated. Overlaid on the city is a network of white lines connecting various points, with several Wi-Fi signal icons (three curved lines) placed at the nodes of the network, suggesting a smart city or IoT environment.

Boost Performance of Financial Services IoT Projects with In-Memory Computing

A GridGain Systems In-Memory Computing White Paper

July 2017

Contents

In-Memory Computing: The Right Choice Now	2
Apache® Ignite™ Tackles Huge Data Volumes	3
FinTech Developers Can Take Advantage of In-Memory Benefits	3
Deploying In-Memory Computing for Optimal IoT Implementations	5
In-Memory Computing Supports Mobile Banking	5
In-Memory Computing for IoT Platforms	6
GridGain Systems: A Leader in In-Memory Computing	6
IoT Will Mean Money in the Bank	7
Contact GridGain Systems	8
About GridGain Systems	8

Beyond all the buzzwords, the Internet of Things (IoT) is shaping up to be very big business. Imagine 26 billion devices—not counting computers and smartphones—creating data, communicating with the Internet, and resulting in \$1.9 trillion in global economic value-add. Those numbers, reported by Gartner, echo a report from IDC which calculated \$737 billion in spending on IoT equipment and services in 2016. Add to that an often-cited EMC prediction that by 2020, the digital universe will contain 44 trillion gigabytes of data, and the data management challenges of the future become readily apparent.

The financial services industry is already embracing IoT devices, collecting data that needs to be analyzed in real time and stored for historical analysis. Car insurance providers use telematics to monitor driver behavior, while health insurance providers are eager to connect with fitness tracking devices. Banks are installing location-aware ATMs that pre-load account information from approaching customers who are identified by their cell phones. Could an investment manager get a health alert from a client's wearable monitor and readjust the client's portfolio to limit risk as an emergency measure? All sorts of ideas are moving from the realm of the possible to the probable.

The platform for all this data collection, storage, and analysis must have several winning characteristics:

- Highly efficient sensors and devices
- Ubiquitous high-bandwidth network connectivity
- High availability
- Fast and scalable back-end storage, computational, and analytical systems
- Streaming data collection in near real time
- The best possible security
- The ability to adjust to variable workloads
- Interoperability

As IoT devices become part of our daily lives and generate huge amounts of streaming data, businesses are going to have to find a way to cope with the constant influx of data, and financial services companies will be on the leading edge. It is time for them to look at in-memory computing solutions.

In-Memory Computing: The Right Choice Now

In-memory technology has been around for decades, ever since developers realized that working with data stored in memory was faster than working with data stored on disk. Today, that speed means everything. Digital transformation has pushed companies closer to customers who demand real-time interactions, and the amounts of Internet traffic, connected devices, and data—all of which require instant analysis—continue to grow.

Luckily, the cost of memory continues to tumble—down 30 percent annually over time—so moving to in-memory platforms has gotten more cost-effective with each passing year. As a result, Gartner has reported, the in-memory technology market will grow to \$10 billion by the end of 2019, representing 22 percent compound annual growth.

The combination of these two trends means that the time is right for the financial services industry to leverage the improved performance and scale provided by in-memory computing to turn their IoT plans into action.

Apache® Ignite™ Tackles Huge Data Volumes

According to Gartner, the in-memory computing market includes in-memory databases, in-memory analytics, in-memory data grids, in-memory app servers, in-memory messaging, event stream processing, and analytic servers. The GridGain and Apache Ignite in-memory computing platforms include many of these components.

Apache Ignite:

- Supports data caching, massive parallel processing, distributed SQL, streaming, and much more
- Slides in between the existing application and data layers easily, with minimal changes to code
- Offers ACID compliant transactions
- Works with all popular RDBMS, NoSQL, and Hadoop databases and offers a unified API with support for a wide range of languages
- Can be used as a memory-centric, distributed SQL database
- Can be deployed on premise, in the cloud, or in hybrid environments
- Can be configured for high availability with automatic load balancing and failover
- Remains highly customizable for a multitude of uses as it handles computing and transacting on large-scale data sets in real time

The GridGain and Apache Ignite in-memory computing platforms can process the full dataset in memory to yield the performance and scalability gains anyone managing huge IoT data streams will need in a hybrid transactional/analytical processing (HTAP) environment. The GridGain Enterprise Edition adds beneficial enterprise features such as enhanced security, network segmentation protection, data center replication, and rolling updates that allow sequentially updating the nodes of a cluster without taking the whole cluster down. GridGain and Apache Ignite can also be deployed using the Persistent Store feature as memory-centric, distributed, transactional SQL databases which hold the complete dataset on disk and all or a subset of the dataset in memory while allowing data from both disk and memory to be processed. The GridGain Ultimate Edition offers a Cluster Snapshot capability which provides backup capabilities to support deploying the Persistent Store feature in production environments.

FinTech Developers Can Take Advantage of In-Memory Benefits

GridGain Systems has a wide range of customers in financial services industries including banking, insurance, and investment management. In a 2016 survey, GridGain asked them what use cases they were considering for in-memory computing. The top uses: database caching, application scaling, high-speed transactions, and real-time streaming, all of which would serve an IoT project well. GridGain Systems asked them where they would run such a system, and the two leading responses were on

premise and in a private cloud followed by deploying on AWS, Microsoft Azure, and in another public cloud. The flexibility of GridGain and Ignite lets them run anywhere.

Asked which languages the respondents use to manage their data, Java and SQL were the two leaders by far. C++, MapReduce, Scala, and .NET followed. It is interesting to see C++ rank this high as it is a language that has earned favor in the financial services community.

Asked which data stores they they were likely to use as their persistent layer underneath GridGain and Apache Ignite, respondents ranked Oracle, MySQL, Microsoft SQL Server, and HDFS in order.

Survey respondents also said that Data Grid, Compute Grid, and the In-Memory File System were the three most compelling features of the GridGain system. This response makes sense given the importance of data scaling and the need for distributed computing platforms in financial services.

Digging deeper into GridGain's in-memory computing platform, we learn how those components work together. The powerful starting point is the **In-Memory Data Grid**, inserted between the application and database layers to cache the disk-based data from the RDBMS, NoSQL, or Hadoop databases in RAM, and deliver a performance increase up to 1,000 times. This in-memory key value store can replicate and partition data caches across multiple nodes and deliver elastic on-demand scalability to add new nodes. Distributed in-memory transactions can also be ACID-compliant. The Data Grid offers support for all popular RDBMSs, with read-through and write-through and support for write behind. Setup is flexible to address unique use cases.

The **In-Memory Compute Grid** enables distributed parallel processing of resource-intensive compute tasks. It offers adaptive load balancing, automatic fault tolerance, linear scalability, and custom scheduling. Built around a pluggable SPI design, it offers a direct API for Fork-Join and MapReduce processing.

The **Distributed SQL** is horizontally-scalable, fault-tolerant, and ANSI SQL-99 compliant with support for all SQL, DML and DDL commands such as SELECT, UPDATE, INSERT, MERGE, and DELETE queries or CREATE or DROP tables. It is a mature, in-memory solution to supplement or replace a disk-based RDBMS. Geospatial support is built into the product and all the communication to the SQL grid is done through ODBC and JDBC APIs without custom coding.

Persistent Store is a distributed ACID and ANSI-99 SQL-compliant disk store available in Apache Ignite that transparently integrates with GridGain as an optional disk layer. It may be deployed on spinning disks, solid state drives (SSDs), Flash, 3D XPoint or other similar storage technologies. Persistent Store keeps the full dataset on disk while putting only user-defined, time-sensitive data in memory. With Persistent Store enabled, users do not need to keep all active data in memory or warm up their RAM following a cluster restart to utilize the system's in-memory computing capabilities. The Persistent Store keeps the superset of data and all the SQL indexes on disk, which allows GridGain to be fully operational from disk. The combination of this feature and the platform's advanced SQL capabilities allows GridGain to serve as a distributed transactional SQL database, spanning both memory and disk. Persistent Store allows organizations to maximize their return on investment by establishing the optimal tradeoff

between infrastructure costs and application performance by adjusting the amount of data that is kept in-memory.

The **In-Memory Service Grid** provides control over services deployed on each cluster node and guarantees continuous availability of all deployed services in case of node failures. It can automatically deploy services on node startup, deploy multiple instances of a service, and terminate any deployed service. It is a load-balanced and fault-tolerant way of running and managing services across the grid.

In-memory streaming and continuous event processing establish windows for processing and run either one-time or continuous queries against these windows. The event workflow is customizable and often used for real-time analytics. Data can be indexed as it is being streamed to make it possible to run extremely fast distributed SQL queries against the streaming data.

In-memory Hadoop acceleration provides easy-to-use extensions to disk-based HDFS and traditional MapReduce, delivering up to 10 times faster performance. GridGain and/or Ignite can be layered on top of an existing disk-based HDFS and used as a caching layer offering read-through and write-through while the GridGain Compute Grid can run in-memory MapReduce.

In other words, when streaming data floods in from IoT sources, users can store it in the Data Grid, and work with it with machine learning and all kinds of real-time analytics using the distributed computing architecture of Compute Grid. In-Memory Streaming and Complex Event Processing comes into play as users manage what is an unending data stream, setting event windows based on time or a specific volume of data and running operations or analytics even as the data continues to flow.

With all these features and tools enabled and running optimally, financial services developers can vastly improve their applications by:

- Processing high-speed transactions faster than ever
- Speeding up search dramatically
- Deploying better systems for the management of IoT components

Deploying In-Memory Computing for Optimal IoT Implementations

In-Memory Computing Supports Mobile Banking

Many financial services firms have adopted GridGain's in-memory computing solution to improve performance and functionality. Clients such as Apollo Global Management, Barclays, Cambridge Associates, Citi, Fidelity, and ING have taken advantage of GridGain's distributed computation and in-memory computing speed.

Sberbank, the largest bank in Eastern Europe and the third largest in Europe with over 16,000 branches across 11 time zones, was experiencing significant growth in its transaction volume. Driven by mobile banking, transactions had increased from 30 per second a decade ago to 3,000-4,000 per second today. Sberbank wanted to eliminate its vertical software stacks for business lines like payments and

mortgages and minimize its legacy dependency on Oracle by migrating to an open source data grid architecture for its next generation platform. The bank needed to introduce new products in hours, not weeks, and the platform had to have virtually unlimited performance and very high reliability while lowering costs and reducing human interaction.

The GridGain solution Sberbank implemented offers a very high level of redundancy and high availability and has the security levels Russian regulators demand. In testing, the bank was able to execute up to one billion transactions per second on an array of ten Dell R610 blades with one terabyte of memory, assembled at a cost of just \$25,000.

In-Memory Computing for IoT Platforms

GridGain Systems worked with Silver Spring Networks, a company which provides an IoT platform for Internet-connected devices such as power meters and other smart cities devices, to address a large real-time data processing challenge. Silver Spring Networks connected energy meters deliver data from millions of homes and the company forecasted that its existing database infrastructure would soon no longer be able to keep up with the growth of its connected devices. Not only did the Silver Spring Networks IoT platform need to ingest massive amounts of data constantly, it also had to overlay current electricity rates and perform compute-heavy calculations on the fly to produce regular analytical reports combining the current state of the power grid with its recent history.

Silver Spring Networks was committed to keeping as much of its database infrastructure intact as possible, so the solution was to insert the GridGain in-memory computing platform between the existing data and application layers. This resulted in a high performance, distributed computing environment in which to do the required real-time data processing and analytics. The system provides as much scalability as the company needs allowing it to grow to thousands of nodes, if necessary, as its business grows. Today the company can store a week of its current data entirely in RAM, vastly improving its ability to deal with the never-ending incoming data it must process in real-time.

GridGain Systems: A Leader in In-Memory Computing

The GridGain in-memory computing platform helps users consolidate onto a single high-performance and highly scalable solution for transacting and analyzing large datasets in real-time, resulting in lower TCO. Advanced SQL functionality and API-based support for common programming languages enable rapid deployment. Coupled with the rapidly decreasing cost of memory, this boosts ROI for in-memory computing initiatives, enabling users to create solutions that can perform thousands of times better. Any financial services company wrestling with large data volumes and seeking to deploy impactful solutions can realize the benefits of the GridGain in-memory computing platform.

Clients enjoy the following:

- **A unified high-performance architecture.** The GridGain in-memory computing platform consists of multiple grids connected by a clustered in-memory file system. The In-memory Data Grid, Compute

Grid, Distributed SQL, Streaming Grid and Service Grid are interconnected. Computations occur as close as possible to the data used in the computation. Additional features such as high throughput, low latency, load balancing, caching, in-memory indexing, streaming, Hadoop acceleration and other performance improvements are crucial to success in real-time modeling, processing, and analytics.

- **Scalability.** The GridGain in-memory computing platform excels in terms of scalability, allowing users to add cluster nodes and memory in real-time with automatic data rebalancing. As a hardware-agnostic solution, clients can choose their preferred hardware for scaling up.
- **Full SQL support.** GridGain is ANSI SQL-99 compliant and the Distributed SQL capabilities include DDL and DML support which allows users to leverage their existing SQL code using the GridGain JDBC and ODBC APIs. For users with existing code bases which are not based on SQL, they can leverage their existing code through supported APIs for Java, .NET, C++, and more.
- **High availability.** The GridGain in-memory computing platform is a distributed computing environment which is highly available and includes related high availability features such as data-center replication, automatic failover, fault tolerance, and quick recovery on an enterprise-level scale.
- **ACID transactions.** The GridGain in-memory computing platform supports ACID compliant transactions in a number of user-configurable modes.
- **A cost effective, memory-centric distributed SQL database option.** Using the Persistent Store feature, users can deploy a memory-centric distributed transactional SQL database which stores the full dataset in the GridGain disk store and holds a subset of the data in memory. Transactions and analytics can be completed on the entire dataset, whether the required data is in memory on disk. By adjusting the amount of system memory, users can optimize the cost versus performance of their system.
- **Security features.** The GridGain in-memory computing platform supports authentication, authorization, multiple encryption levels, tracing, and auditing.
- **Open source framework.** GridGain is based on Apache® Ignite™, a popular open source project with many contributors that has been tested globally. GridGain Systems was the original creator of the code contributed to the Apache Software Foundation that became Apache Ignite and fully supports the technology behind Apache Ignite. The [GridGain Enterprise Edition](#) extends the feature set in Apache Ignite with enterprise-level capabilities such as enhanced security, data center replication, a GUI for management and monitoring, and network segmentation protection.
- **Production support.** [GridGain Systems Support](#), available to [GridGain Ultimate Edition](#), [GridGain Professional Edition](#) and [GridGain Enterprise Edition](#) users, includes rolling updates, faster availability of all releases and patches, and 24/7 enterprise-level support.

IoT Will Mean Money in the Bank

In-memory computing gives financial services companies the speed, scalability, and flexibility they need to build IoT-based applications and services. This allows them to serve their current clients in new and profitable ways while attracting new clients as well. No matter what kinds of Internet of Things implementations come to fruition in the future, the high-performance transactions and real-time analytics across a variety of platforms that in-memory computing offers will be the foundation on which they are built.

Contact GridGain Systems

To learn more about how the GridGain in-memory computing platform can help your business, please email our sales team at sales@gridgain.com, call us at +1 (650) 241-2281 (US) or +44 (0) 7775 835 770 (Europe), or complete our [contact form](#) to have us 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, ecommerce, retail, online business services, healthcare, telecom and other major sectors, with a client list that includes Barclays, ING, Sberbank, Misys, IHS Markit, Workday, Silver Spring Networks and Huawei. GridGain solutions can connect data stores (RDBMS, NoSQL and Apache® Hadoop®) with web-scale applications or can function as a standalone transactional SQL database to enable massive data throughput and ultra-low latencies across a scalable, distributed cluster of commodity servers. GridGain is the most comprehensive, memory-centric in-memory computing platform for high volume ACID transactions, real-time analytics, and hybrid transactional/analytical processing. For more information, visit gridgain.com

COPYRIGHT AND TRADEMARK INFORMATION

© 2017 GridGain Systems. All rights reserved. This document is provided “as is”. Information and views expressed in this document, including URL and other web site references, may change without notice. This document does not provide you with any legal rights to any intellectual property in any GridGain product. You may copy and use this document for your internal reference purposes. GridGain is a trademark or registered trademark of GridGain Systems, Inc. Windows, .NET and C# are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries. JEE and Java are either registered trademarks or trademarks of SUN Microsystems and/or Oracle Corporation in the United States and/or other countries. Apache, Apache Ignite, Ignite, the Apache Ignite logo, Apache Spark, Spark, Apache Cassandra, and Cassandra are either registered trademarks or trademarks of the Apache Software Foundation in the United States and/or other countries. All other trademarks and trade names are the property of their respective owners and used here for identification purposes only.