MUTABLE AWARD 2019
GOLD

GridGain

www.GridGain.com

1065 East Hillsdale Blvd. Suite 220
Foster City, CA 94404, USA
Tel: +1 650 241 2281
Email: info@GridGain.com

GridGain Systems

The company

GridGain was formed initially in 2007 but not incorporated until 2011. In 2014 it donated its code base to Apache Software Foundation as the open source Apache Ignite project. The company now (April 2019) provides support for Ignite as well as offering Professional, Enterprise and Ultimate Editions that provide additional capabilities over and above those of Ignite. The products are available for deployment on-premises, in the cloud, or in hybrid environments, and they are available as a GridGain Cloud managed service as well as through the AWS, Microsoft Azure and Oracle Cloud Marketplaces.

> GridGain has enabled us to provide a high performance, highly responsive, and highly available hosted platform for our partners. GridGain is extremely flexible, works as advertised, and is vital to meeting our strict SLAs.
>
> FSB Technology (UK)

GridGain is funded through VC backers and is based in California but also has offices in London, Paris, Russia and Hong Kong. The company has an extensive partner program, across systems integrators, cloud platforms, independent software vendors and technology providers. In the last category, perhaps Confluent (Kafka) is
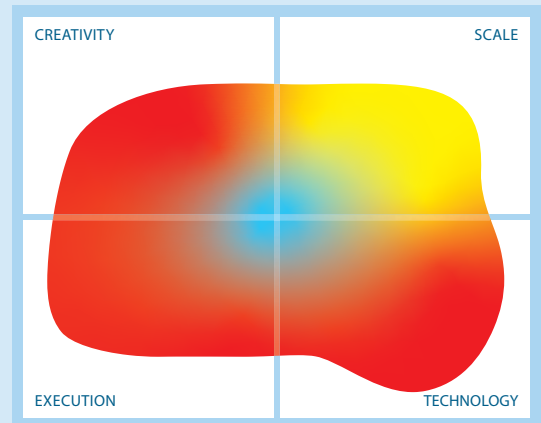
The image in this Mutable Quadrant is derived from 13 high level metrics, the more the image covers a section the better. Execution metrics relate to the company, Technology to the product, Creativity to both technical and business innovation and Scale covers the potential business and market impact.

the most significant. GridGain initially focused on the financial sector but now has customers spanning industry sectors and including organisations such as American Express, ING, Marketo, Tui and American Airlines.

What is it?

GridGain and Ignite started life as in-memory data grids. However, unlike most other data grids, both products have now evolved to include the option to act as an in-memory database (see *Figure 1*) which means that they are now, technically, in-memory computing (IMC) platforms rather than just data grids. When the database option is chosen, data is persisted to a disk-based key-value store. It also enables GridGain to process data from both memory or disk, with processing on the data on disk either because there is insufficient memory or because this is *"colder"* data that will not normally be available in memory. Having said all of that, most deployments continue to be as an in-memory, distributed data grid that front-ends onto – most often – a traditional relational database, bringing the performance and horizontal scalability that would otherwise be very costly to implement, and without requiring any rip and replace strategy.
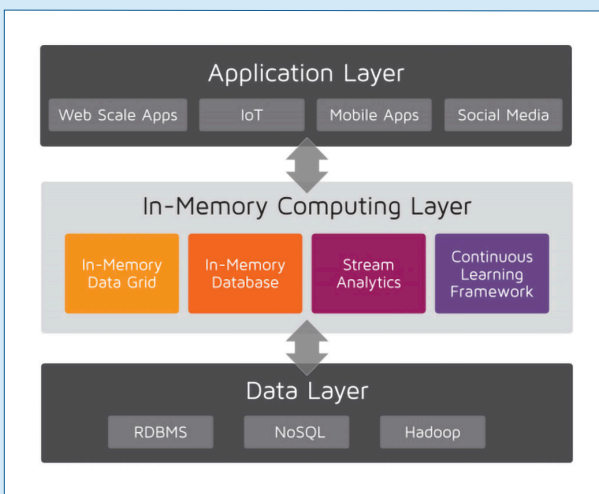

Figure 1 – Showing the in-memory computing platform

Architecture	★★★★⯪	Flexibility	★★★★
Concurrent analytics	★★★★⯪	In-process analytics	★★★★⯪
Criticality	★★★★⯪	Performance	★★★★⯪
Ease of use	★★★★⯪	Scale	★★★★⯪

How does it work?

GridGain uses a shared nothing architecture with data nodes that hold both data and indexes and client nodes that act as connection points. The former process incoming requests along with computations. Partitioning is supported with automatic cluster rebalancing whenever a node is added or removed. Support is provided for ODBC, JDBC and REST and a wide variety of languages are supported, including ANSI standard SQL. Transaction processing comes with ACID guarantees while strong consistency is provided along with support for two-phase commit.

Figure 2 – Illustrating the Continuous Learning Framework

In so far as *Figure 1* is concerned, GridGain enables the ingestion of streaming data through a Kafka connector, and there is also support for both Spark DataFrames and Spark Resilient Distributed Datasets (RDDs). Perhaps most interesting, however, is the Continuous Learning Framework (CLF). This is illustrated in *Figure 2*. The point is that machine learning and other analytic models can be trained and tested in the platform, which means that those models can be updated directly rather than periodically, which tends to be the case when models have to be trained and tested in an outside environment. This also lends itself to environments where you want to embed analytics *"in process"* within an otherwise transactional or operational application. While it will often be the case that you

want to run these analytics within the GridGain environment, this does not have to be so. You can, for example export models if, for example, you wanted to run these on GPUs for extra processing performance.

While we have already mentioned support for Spark (and Spark ML), GridGain is also integrated with TensorFlow for deep learning. While the company does not specifically support PMML (predictive modelling mark-up language) it can leverage the PMML support provided by TensorFlow or Spark though it would not help if you wanted to import a model derived using SAS, as an example.

Why should you care?

The big benefit of using a data grid is that you can simply add it to your existing environment without any of the pain of replacing your current database. On the other hand, there are times when you positively want to replace existing systems and, of course, there are greenfield developments where you need a new database from the get-go. The optional deployment of GridGain as an in-memory database, as opposed to a pure data grid, offers the flexibility to support all of these scenarios.

On top of that, we are impressed with the Continuous Learning Framework. In hybrid environments where you want to support concurrent transactional and analytic processing, there are two generic use cases: analytics (typically dashboards) that run in parallel to operational functions and analytics that are embedded (in-process) in operational applications. Everybody in this space supports the former but GridGain is one of relatively few vendors that offer specific capabilities to support the latter.
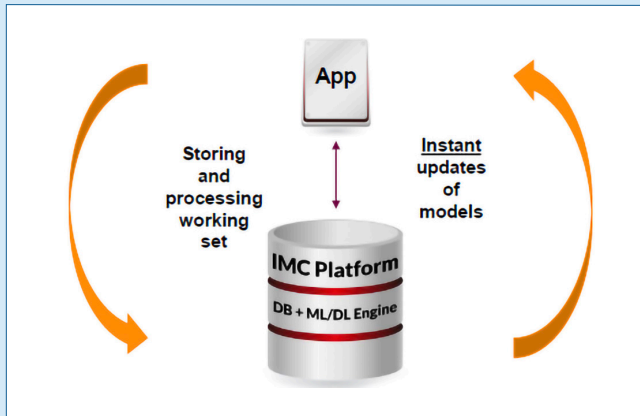
The Bottom Line

There is lots of competition in the market for in-memory databases but that is much less true for in-memory data grids. In the latter case, Apache Ignite is by far the most widely deployed offering and GridGain is thus positioned as the market leader in this category: if you want to augment your existing environment with in-memory technology GridGain is an obvious possible choice.

FOR FURTHER INFORMATION AND RESEARCH CLICK HERE