insideBIGDATA

InsideBIGDATA Guide to

In-Memory Computing

by Daniel D. Gutierrez

BROUGHT TO YOU BY GridGain

Introduction to In-Memory Computing

In-memory computing (IMC) is an emerging field of importance in the big data industry. It is a quickly evolving technology, seen by many as an effective way to address the proverbial 3 V's of big data — volume, velocity, and variety. Big data requires ever more powerful means to process and analyze growing stores of data, being collected at more rapid rates, and with increasing diversity in the types of data being sought — both structured and unstructured. In-memory computing's rapid rise in the marketplace has the big data community on alert. In fact, Gartner picked in-memory computing as one of the Top Ten Strategic Initiatives.

> Put simply, in-memory computing primarily relies on keeping data in a server's RAM instead of much slower spinning disk or flash devices and massive parallelization as a means of processing at faster speeds.

But what is In-memory computing? Put simply, in-memory computing primarily relies on keeping data in a server's RAM instead of much slower spinning disk or flash devices and massive parallelization as a means of processing at faster speeds. In-memory computing especially applies to processing problems that require extensive access to data analytics, reporting, data warehousing, high-speed transactions and big data applications.

There are two primary ways for how IMC solves the problems of big data — lowering latency time for analytics applications and yielding higher throughput for transactional applications. By removing high-latency devices like disk drives out of the equation, the time it takes to process a request and deliver a response (latency) is dramatically reduced. Estimates vary depending on disk speed and available input/output (I/O) bandwidth, but by one measure, RAM latency is estimated at 83 nanoseconds and disk latency at 13 milliseconds, a whopping speedup of 6 orders of magnitude. You can't own this entire speed advantage because there are CPU processing time and other constraints in the mix, but disk I/O has long throttled performance. In-memory performance improvements vary by application, data volume, data complexity, and concurrent-user loads, but by any measure the speedup with IMC can be dramatic.

The case for IMC and analytics applications is clear, but what about mainstream transactional systems? IMC represents a significant opportunity when you can run a query, process, or transaction 10x, 20x, even 100x faster than you used to. What will that difference do for your business? The answer will determine whether IMC is a specialty tool for only a few unique use cases, or a platform upon which your entire enterprise IT runs. The purpose of this Guide is to provide the information that enterprise thought-leaders need to make strategic decisions about the use of IMC.

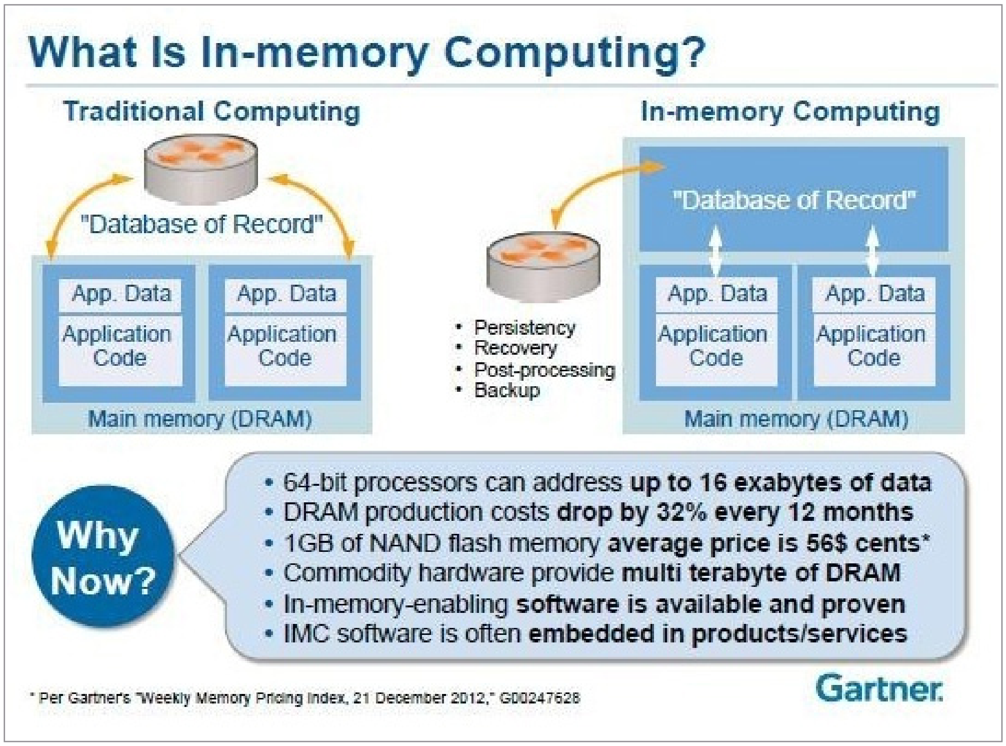
Contents

Let's take a step back here and review what is meant by "in-memory." You often see discussions of IMC with references to DRAM (dynamic random-access memory), SSD (solid state disk), PCI-based flash (server-side flash), memory channel storage (non-volatile DIMM memory that delivers low-latency storage directly on the processor bus), and others. Most storage vendors provide a tiered storage model where part of the data is stored in DRAM which is complemented by a variety of flash or disk devices. Rarely is there a DRAM-only, Flash-only, or disk-only product. Generally speaking however, it is DRAM that best deserves the label as facilitator for IMC.

The most important factor for why IMC is a reality now: DRAM costs are dropping about 32% every 12 months. Big data is getting bigger, and costs are getting lower. If you looked at the price of a Dell server with a terabyte of memory three years ago, it was almost $100,000. Today, a server with more cores — sixteen instead of twelve — and a terabyte of DRAM, costs less than $40,000.

The high-level architecture of a traditional and in-memory data warehouse architecture is shown in the figure below. The shift from traditional, hard disk-based data warehouses to IMC-based data warehouses yields an organization with fewer layers along the path from raw data to analytics results.

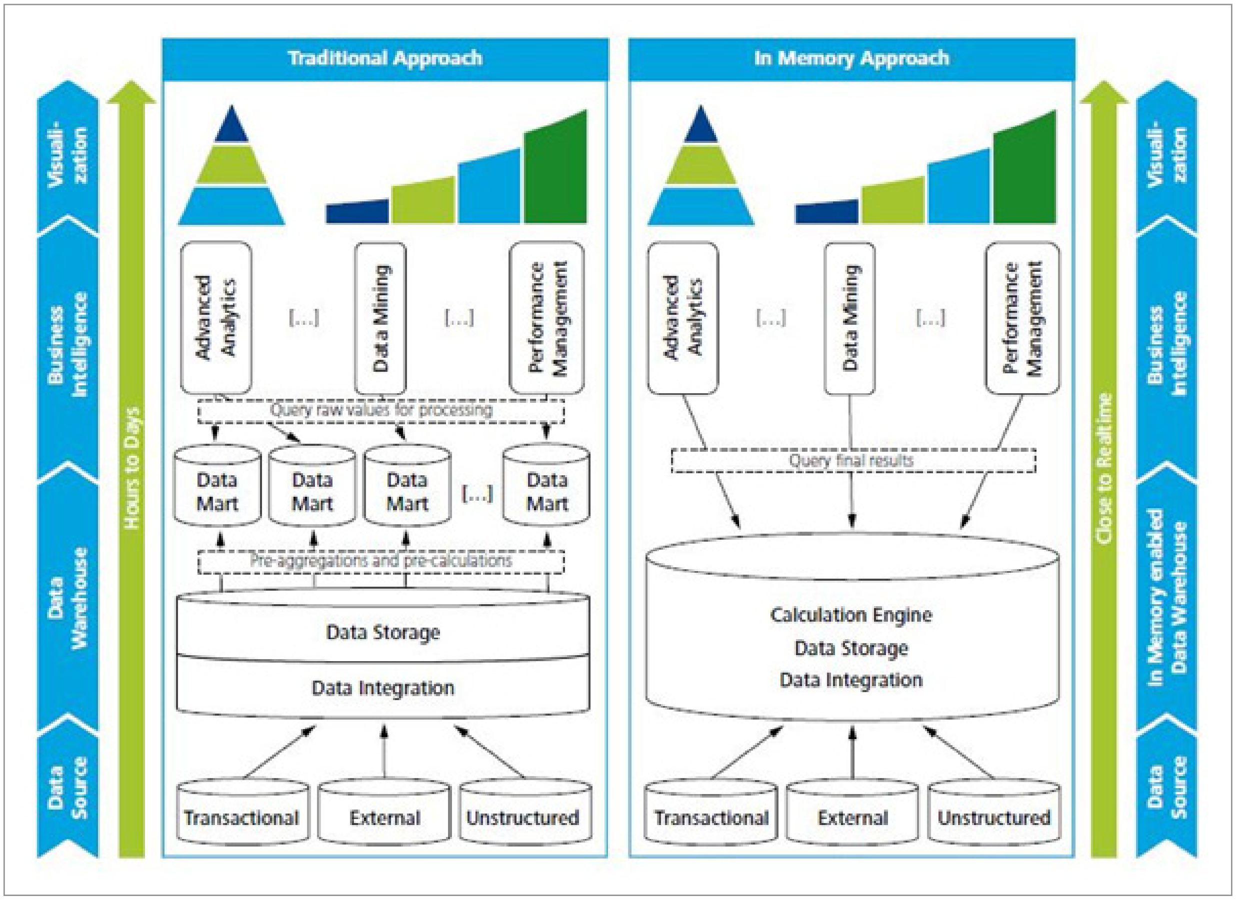
With traditional data warehouses, raw data are stored in a data warehouse and portions of the data are extracted to data marts for use by specific departments — providing user specific aggregations and calculations. The content stored in a data mart is then utilized by business intelligence applications for the processing and visualization of final results. As in a traditional data warehouse, raw data is stored in an IMC-enabled data warehouse, however business intelligence applications do not request partial results from data marts but rather final results from the data warehouse as a whole. The data mart layer becomes obsolete. Furthermore, IMC-enabled data warehouses allow for a frequent update of raw data so that transactional applications can directly feed data into an IMC-based data warehouse.



A Gartner report summarizes IMC and why the technology is now a reality, in particular the drop in DRAM costs.

IMC is becoming a favorite approach in transactional systems such as ERP or financial trading applications with the expectation that IMC will both improve performance for line-of-business applications and allow more sophisticated analytics within them. This means that data-intensive processes and transactions can be performed much faster.

As with many new fields of technology, the genesis often arises from academic research. In the case of IMC, the technological foundation comes from the mid-1980s. At the time, so-called "database machines" received a hotbed of research activity in an effort to build computer architectures specifically designed for database operations. An early research paper that defined initial ideas about IMC was *MARS: The Design of a Main Memory Database Machine*[1]. At that time, the goal was to build a machine with an address space of 1-20 GB which was very large for that period.



Traditional data warehouse approach vs. IMC-enabled data warehouse approach from "In-Memory Computing Technology – The Holy Grail of Analytics?" a research report by Deloitte & Touche

[1] Margaret H. Eich, "MARS: The Design of a Main Memory Database Machine," Proceedings of the 5th International Workshop on Database Machines, Karuizawa, Japan, October 1987, pp 468-481.

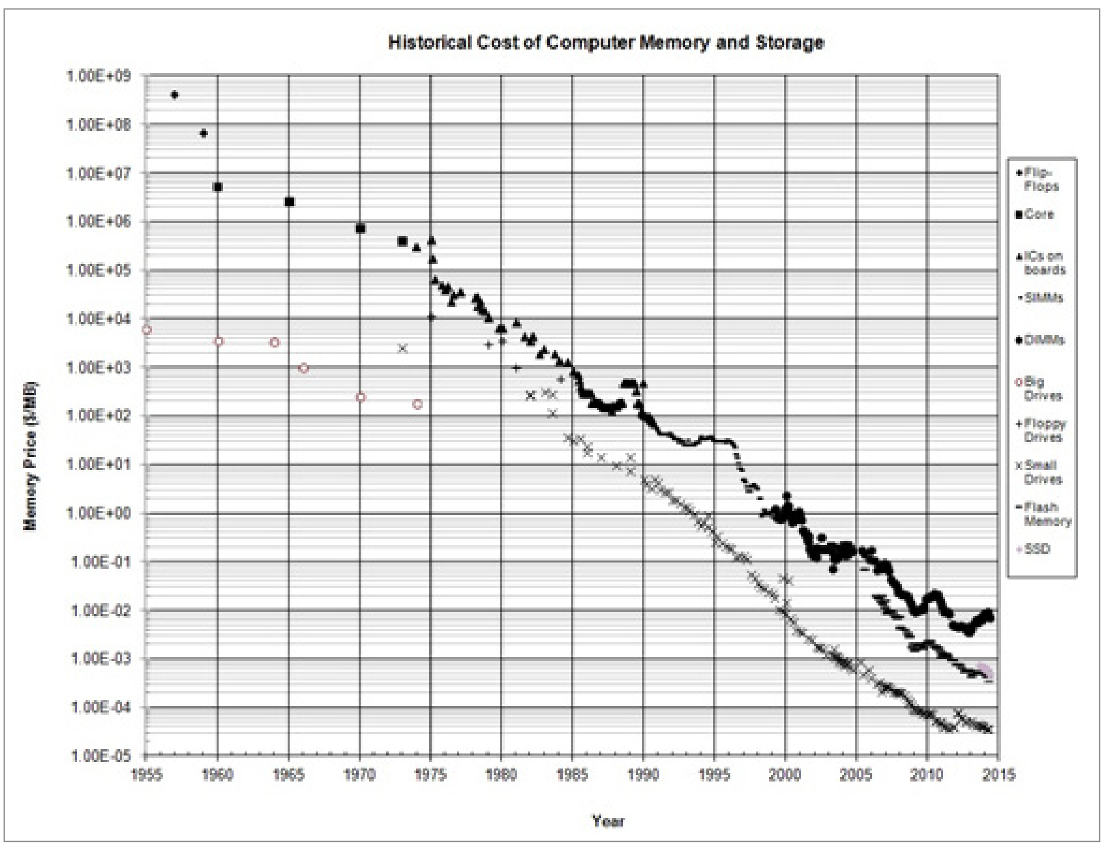
The Business Case for In-Memory Computing

A confluence of factors have contributed to the strong business case for IMC, including the steep upward slope for the collection of business data, the demand for analytics approaching real-time, and increased complexity of business software applications. In light of these trends, added performance translates directly into clear business value:

- With IMC you can use less hardware to (i) support the required SLAs directing performance and throughput, (ii) get better data center consolidation, and (iii) significantly reduce capital costs, as well as infrastructure and operational overhead.
- With IMC you can significantly extend the lifetime of your existing hardware and software by getting increased performance and improve its ROI by using your current infrastructure longer and making it perform better.

And that's what makes IMC such a hot topic these days: the demand to process ever growing data sets in real-time can now be fulfilled with the extraordinary performance and scale that IMC affords. Enterprise-ready solutions using IMC technology can help companies analyze data faster, improve the quality of their business decisions, and use the insight to increase customer satisfaction and sales performance. With the economies-of-scale equation so compelling, the business case becomes clear and obvious.

Gartner reported in 2013 that IMC is racing towards mainstream adoption. "The relentless declines in DRAM and NAND flash memory prices, the advent of solid-state drive technology and the maturation of specific software platforms have enabled IMC to become more affordable and impactful for IT organizations," said Massimo Pezzini, vice president and Gartner Fellow. Since 1995, DRAM and NAND flash memory prices have plunged from approximately $30/MB to $.0073/MB.



Graph of memory prices decreasing with time (1957 – 2014) courtesy of John C. McCallum.

IMC is an inexpensive means for speeding up enterprise software applications, including but not limited to analytics. The question, however, is what do we mean by speed-up? If we mean a speed gain of 2-3x performance or scalability improvements, then flash storage is the path of least resistance. This route is most inexpensive and provides a modest increase in performance. But what we now think of as IMC, extends well beyond this level of performance gain.

> A company simply cannot maximize the business impact of today's technology without access to both data and performance.

IMC used to be a strategy only for larger enterprises with deep pockets, but now due to the significantly deceased price of RAM, businesses of all sizes can take advantage of this technology to process real-time data. The main advantage of IMC is the storage of data in the RAM on multiple servers, which decreases the impact of the data on performance. By storing data in RAM instead of on a mechanical spinning disk, the performance increases exponentially. Without IMC, companies do not have access to the full capabilities of technologies such as big data. A company simply cannot maximize the business impact of today's technology without access to both data and performance.

As an example of a real-life use case for IMC, consider a client engagement that IMC industry leader GridGain won with one of the largest banks in the world. The solution called for a risk analytics system to provide real-time analysis of risk for the bank's investment portfolio (a common use case for IMC in the financial industry). GridGain software demonstrated 1 billion (full ACID) business transactions per second using 10 commodity servers with a total of 1TB of RAM. The total hardware cost for the 10 commodity servers was less than $25,000. This example demonstrates a dramatic business case for IMC. IMC did not just provide a mere 2-3x performance boost, but rather a staggering 100x of what is theoretically possible using the most expensive flash-based storage available on today's market. 1TB of flash-based storage alone would cost 10x of entire hardware setup mentioned above. Hard disks wouldn't even have been in the same realm of consideration.

IMC's impact is increasingly manifesting as notable innovation in business processes. Well-established industries and application areas already have adopted IMC including: manufacturing resource planning (MRP), corporate performance management, dynamic pricing and supply chain planning (SCP), insurance claim processing and modeling, medical image processing, natural language processing, real-time machine learning, real-time ad platforms, investment banking, and much more. It is clear that these use case examples are driven by the underlying technical need, not by the specific industry. The end goal for any data environment should be delivering precisely the right information to the right people quickly enough for prompt action — IMC certainly facilitates this goal. Longer term, even more innovation is destined to emerge as more IMC-enabled packaged business applications are released in the market.

Types of In-Memory Computing

In this section we'll set the stage for in-memory computing technology in terms of its current state as well as its next stage of evolution. We'll begin with a discussion of the capabilities of in-memory databases (IMDBs) and in-memory data grids (IMDGs), and show how they differ. We'll finish up the section by demonstrating how neither one is sufficient for a company's strategic move to IMC; instead, we will explain why a comprehensive in-memory data platform is needed.

Traditionally, IMC is comprised of two primary classes of technology: in-memory databases (IMDBs) and in-memory data grids (IMDGs). In-memory databases come in two flavors — traditional databases with in-memory options such as Oracle 12c (including Exalytics and Exadata), IBM DB2 with BLU Acceleration, and Microsoft SQL Server 2014. The IMC offerings from traditional RDMS vendors can be considered point solutions (Band-Aid solutions) that represent a single feature in a long list of advertised tools, rather than a more strategic approach toward in-memory architecture. There also are native in-memory databases like Altibase, MemSQL, VoltDB, EXASOL, H20, SAP Hana and others.

IMDGs, on the other hand, provide MPP-based (Massively Parallel Processing) capabilities where data is processed in parallel fashion and dispersed across a large cluster consisting of commodity servers. This organization has been made popular with Hadoop's MapReduce methodology.

With IMDGs and their horizontally scalable MPP architecture, there is an inherent capability to scale to hundreds or even thousands of servers.

One way to compare IMDBs and IMDGs is in the level of scalability made possible by the respective technologies. On the one hand, IMDBs utilize SQL that cannot be effectively performed in a distributed environment. This means IMDBs are unable to scale horizontally and as a result, most existing SQL databases (disk or memory based) are based on vertically scalable Symmetrical Processing architecture. But with IMDGs and their horizontally scalable MPP architecture, there is an inherent capability to scale to hundreds or even thousands of servers.

Another point of comparison between IMDBs and IMDGs involves the prerequisite work needed to employ the technology. With IMDBs, replacing existing databases often is required, unless you choose to use an in-memory option to temporarily boost your database performance. This method will require considerably fewer changes to the application itself due to its continued reliance on SQL. On the other hand, IMDGs always work with an existing database, providing a layer of massively distributed in-memory storage and processing between the database and the application. Most IMDGs are highly integrated with existing databases, and can seamlessly perform read and write operations with the databases when necessary. In these instances, developers must modify the application in order to take advantage of these new capabilities but the existing database remains untouched. Here the application no longer is primarily SQL-based, rather it needs to learn how to use MPP, MapReduce or other scalable data processing techniques.

IMDGs also involve the use of a genre of middleware software allowing data to reside in RAM across a cluster of servers, and to be processed in parallel. This means you can take data sets typically stored in a centralized database and now store them in *connected* RAM across multiple servers. RAM is several orders of magnitude faster than a traditional mechanical spinning disk. Add to the mix native support for parallel processing, and overall performance experiences significant gains. RAM storage and parallel distributed processing form the primary tenets of IMC.

Let's take a step back to understand how IMDG parallelization, and distributed data processing are coupled. Parallel distributed processing capabilities of IMDG are essential. Consider a typical scenario: a typical x86 server (in the 2014 time frame) may have a complement of RAM between 32GB to 256GB. Although this capacity could be considered a significant amount of memory for a single server, it is generally not enough to store many of today's operational datasets that easily measure in terabytes.

In order to circumvent this problem, IMDG software is designed from the ground up to store data in a distributed fashion, where the entire data set is divided across the memory of individual computers, each storing only a portion of the overall data set. Once the data is partitioned, parallel distributed processing becomes essential. As a result, end users of IMC-based applications see dramatic performance and scalability increases.

In the end, there are pros and cons associated with both approaches. Coming from the perspective of building a new application, it is advantageous to go with IMDGs since you get to work with existing databases in your organization where necessary, while benefitting from top level performance and scalability. Alternatively, if you're reconditioning an application, you can use an IMDB under the following conditions:

- You're able to upgrade or replace an existing disk-based RDBMS
- You're unable to make changes to your applications
- Increased speed is necessary, but scalability isn't

With the IMDB option, enterprises are able to boost application speed by replacing or upgrading RDBMS without having to adjust the application itself. Conversely, you can use an IMDG under the following conditions:

- Replacing an existing disk-based RDBMS isn't possible
- You're able to make changes to the data access portion of your application
- Both increased speed and scalability are desirable

The following table summarizes the IMDG/IMDB decision making process.

	In-Memory Data Grid	In-Memory Database
Existing Application	Changed	Unchanged
Existing RDBMS	Unchanged	Changed or Replaced
Speed	Yes	Yes
Max. Scalability	Yes	No

Let's finish up this section by reviewing the path traversed by IMC over the recent past and justify the next stage in its development. In-memory computing within the last 5 years, can be characterized by a variety of "point solutions" — IMDGs, IMDBs, in-memory analytics, in-memory business intelligence tools, etc.

As an analogy, look back at the web 20 years ago, and we see plenty of point solutions — different portions of a web server, or application server collectively addressed business requirements. Today, in contrast, we have thoroughly complex web services showing there was an evolution from a quilt work of specific point solutions to more comprehensive offerings. The same thing happened with the cloud as these commonly used technologies all went through the same growth pattern. Companies started to realize the strong benefit of using more cohesive platforms.

If a holistic view of this progress is taken, many CIOs are starting to think of IMC more strategically, and that's the primary motivation for a new generation of IMC technology— the in-memory data fabric.

IMC is in about the same place right now — graduating from a zoo of point solutions to a comprehensive in-memory data platform, or rather an in-memory "data fabric", which is what GridGain calls it. The drivers behind this evolution are very familiar. Customers were giving the same message in unison — "We've been using IMC for the last 5 years, but we don't want half a dozen products in our data center. Rather, we want a more comprehensive and strategic view of IMC. "

To more fully complete the IMC picture, another significant trend is developing — RAM-based data centers where memory will be the primary data store and disks, whether spinning disks or flash, will be used predominantly as back-up devices. What happened to magnetic tapes is what's happening to disk/flash memory. The data centers will become completely RAM-based, especially with non-volatile RAM so if you unplug the power, the data is still there. Further, the speed will be almost identical to DDR3 memory today. This is a huge step forward in the evolution of high-performance, low-carbon-footprint data centers.

If a holistic view of this progress is taken, many CIOs are starting to think of IMC more strategically, and that's the primary motivation for a new generation of IMC technology — the in-memory data fabric. We'll see in the last section of this Guide how GridGain has taken this important step forward in the marketplace.

Performance Benchmark

It is becoming increasingly clear that the key to increasing the speed of processing of both transactional and analytical data, especially for growing workloads and data volumes, is to adopt the IMC paradigm. Building on fundamentally the same parallel computing architecture as supercomputers used in scientific applications, IMDGs run on a clustered set of servers to hold and analyze memory-based data. IMDGs keep access times constant, which is exactly the characteristic needed by applications which have to handle growing workloads. More significantly, IMDGs can host parallelized applications to manage and analyze data stored on the grid's servers. This is the key to their ability to perform real-time analytics.

More significantly, IMDGs can host parallelized applications to manage and analyze data stored on the grid's servers. This is the key to their ability to perform real-time analytics.

The practical benefits of IMC were examined in a research study conducted by Aberdeen Group — "In-Memory Computing: Lifting the Burden of Big Data."[2] The report compared companies that had fully implemented IMC solutions with companies that had not. 196 organizations world-wide, currently engaged with big data, were examined. These companies were chosen to represent a broad cross-section of industries and sizes. The companies dealt with data stores ranging in size from several terabytes to multiple petabytes. A relatively small segment of 33 companies indicated they've implement IMC.

The survey results revealed that companies adopting IMC technology outperformed their peers by two orders of magnitude. They were able to process much more information, must faster and with greater efficiency. Respondents using IMC showed substantially better performance in dealing with the volume of their data. They successfully store large amounts and analyze more data at once, both in terms of size and the percentage of overall business data that could be processed. The real gains, however, came in the velocity category. Organizations with IMC reported an average of 42 seconds for their analytical system to respond to a query. This was over 100x faster than organizations without IMC — these companies saw an average response of an hour and fifteen minutes. The unique ability to analyze data in-memory allowed for rapid processing even with large-scale data sets.

Table 1: More Data, More Speed, More Efficiency

Performance Metrics	Use in-memory computing (n = 33)	Don't use (n = 163)	In-memory Benefit
Median amount of active business data	38 terabytes	18 terabytes	2.1 times more data
Median amount of data analyzed	14 terabytes (37% of all data)	4 terabytes (22% of all data)	3.5 times more data
Average response time for data analysis or query	42 seconds	75 minutes	107 times faster
Data volume processed per hour	1200 terabytes	3.2 terabytes	375 times more efficient

Source – Aberdeen research study "In-memory Computing" (January 2012)

In the table above, the IMC benefit also is determined for the volume of data being managed and analyzed. The impressive results show a clear advantage for IMC where 2.1 times more data is managed (38TB versus 18TB) and 3.5 times more data is analyzed (14TB versus 4TB). These results show that IMC tends to allow companies to manage an ever-increasing amount of data, faster than before.

[2] Nathaniel Rowe, "In-Memory Computing: Enabling Real-time Access to Big Data," Aberdeen Group, March 2013.

GridGain In-Memory Data Fabric

Now that we've taken a tour of the IMC technology and business landscape, let's evaluate a leading open-source product offering by GridGain Systems, Inc. (www.gridgain.com). As we discussed earlier in this Guide, *In-Memory Computing* is characterized by using high-performance, integrated, distributed memory systems to compute and transact on large-scale data sets in real-time, orders of magnitude faster than possible with traditional disk-based or flash technologies. GridGain's solution provides state-of-the-art technology to expertly align with these characterizations and much more. GraidGain was named by Gartner as one of the "Cool Vendors in In-Memory Computing Technologies" for 2014.

The logic for combining a collection of point solutions under a single cohesive platform is clear — they all share the same learning curve, same API, same configuration, same security, same product features. Now all of these important pieces are part of a central fabric.

The technology backdrop for the IMC industry is basically a series of point solutions such as in-memory databases, in-memory data grids, in-memory streaming, etc. With the new GridGain *In-Memory Data Fabric,* these components are all under one central architecture. The logic for combining a collection of point solutions under a single cohesive platform is clear — they all share the same learning curve, same API, same configuration, same security, same product features. Now all of these important pieces are part of a central fabric.

Why is a data fabric an important advantage over individual point solutions? Point solutions are important in the beginning because they address specific customer problems. A complete platform might not be a priority at this early stage. But eventually you have a whole class of related problems, and you don't need a point solution for each one, instead you need a more systematic, fabric kind of approach.

GridGain historically has been a point solution provider. Look back a couple of years, the company had an IMDG product, an IMHPC product, an in-memory streaming product, and recently a Hadoop product. Now the company is moving toward an in-memory data fabric based on all the previously mentioned drivers and benefits. It is a single comprehensive fabric that an organization can adopt to address all different types of data access and data processing problems. It has transactional capabilities to run transactional payloads in-memory in a highly distributed environment. It has analytical capabilities to run analytical payloads, i.e. non-transactional Hadoop based payloads such as MapReduce. You can run traditional HPC payloads. You also have support for streaming payloads that have very different types of algorithms designed to process this type of data.

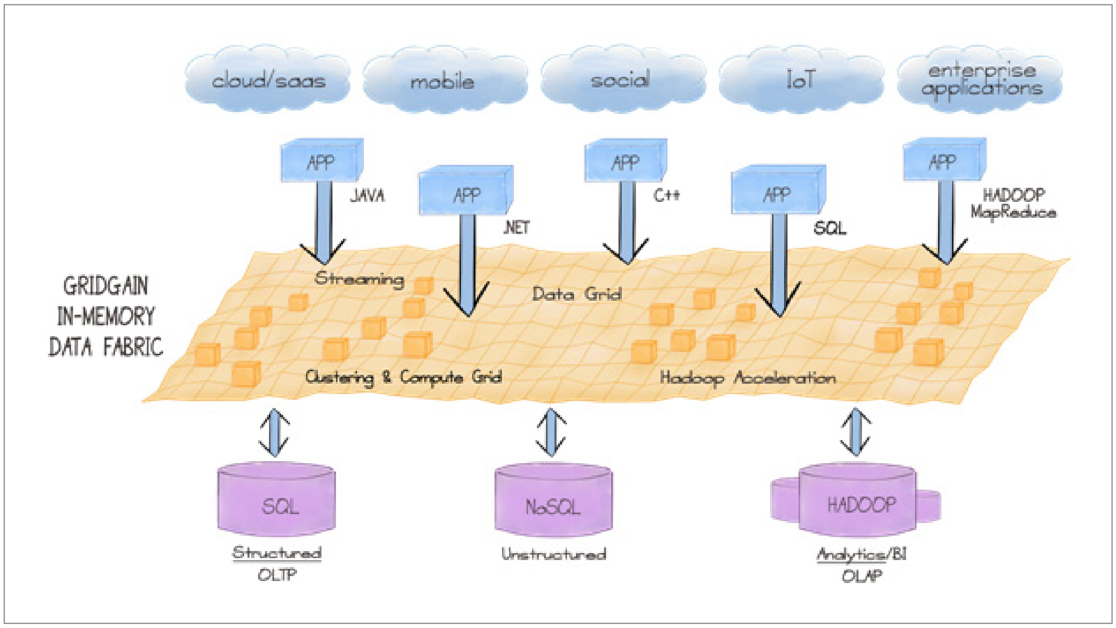
When an organization uses the GridGain In-Memory Data Fabric, it is not limited to one particular type of use case, data type or application. Any type of data processing is possible in-memory with tremendous performance and scalability. This is what a comprehensive in-memory data fabric is all about: Access and process data from any data store — relational, NoSQL, Hadoop — for any application (Java, .NET, C++).

GridGain's IMC solutions are designed to deliver uncompromised performance by providing developers with a comprehensive set of APIs. Developed for the most demanding use cases, including sub-millisecond SLAs, core platform products allow you to programmatically fine-tune performance on large and super-large network topologies with hundreds to thousands of nodes.

The open-source project of the GridGain In-Memory Data Fabric is licensed under Apache 2.0 and is hosted on GitHub where you can review code, learn GridGain internals, and file and review issues. Developers may find it beneficial to look under the hood to understand details, programming style or specifics of implementations.

GridGain understands that In-Memory Computing is more than the latest tech trend. It's the next major paradigm shift for an increasingly data-centric business world in which organizations face problems that traditional technology can't even fathom, much less solve. IMC is a step all organizations must take to remain competitive. The following graphic illustrates the new GridGain In-Memory Data Fabric architecture.

The GridGain In-Memory Data Fabric includes the following unified functional areas: data grid, HPC, streaming, and Hadoop connector. All functional areas were re-engineered and are now fully integrated under the new data fabric architecture—combining all the GridGain technologies into a new cohesive whole. In the following sections, we'll examine each functional area in more detail.



The GridGain In-Memory Data Fabric architecture

Data Fabric Feature: Data Grid

With its In-Memory Data Fabric, GridGain offers industry leading data grid functionality characterized by the fact that data are stored in-memory as opposed to traditional DBMS software that utilizes disk as the primary storage mechanism. By utilizing system memory rather than spinning disk, data grids are typically orders of magnitude faster than traditional DBMS systems. The GridGain data grid feature supports standard SQL for querying in-memory data including support for distributed SQL joins.

GridGain's data grid feature contains an impressive feature set including—Hyper Clustering™, Zero Deployment™, advanced security, fault tolerance, topology resolutions, load balancing, collision resolutions, connected jobs, local node storage and much more. In a clustered in-memory solution like GridGain's, the collection of all individual node memory can be used as a single, expansive "grid" of virtually connected memory. Large data sets can be effectively partitioned across all nodes for high-end scalability, and computations can be intelligently parallelized for optimal processing speed.

At a fundamental level, GridGain enables promoting data up from residing in slow mechanical storage systems to fast memory. GridGain's data grid feature solves many critical IT pain-points at once:

- Performance
- Scalability
- High availability
- Data consistency and reliability
- Detailed insight and management

Data Fabric Feature: Real-time Streaming

To address the needs of a large family of applications for which traditional processing methods and disk-based storages, like databases or file systems, fall short — the GridGain In-Memory Data Fabric offers stream-processing capabilities. In-memory streaming combines both event workflow and *Complex Event Processing* (CEP) capabilities fully integrated in one product.

As streaming data never ends, an application must be able to provide a size limit or a time boundary on how far back each request or each query should go.

Processing of market feeds, electronic trading by many financial companies on Wall Street, security and fraud detection, military data analysis — all these applications produce large amounts of data at very fast rates and require appropriate infrastructure capable of processing data in real-time without blockages.

One of the most common use cases for stream processing is the ability to control and properly pipeline distributed events workflow. As events are coming into the system at high rates, the processing of events is split into multiple stages and each stage has to be properly routed within a cluster for processing.

One of the key features of many CEP systems is the ability to control the scope of operations on streamed data. As streaming data never ends, an application must be able to provide a size limit or a time boundary on how far back each request or each query should go.

Data Fabric Feature: Hadoop Acceleration

With its In-Memory Data Fabric, GridGain offers Hadoop acceleration as well as a standalone In-Memory Accelerator for Hadoop built on top of the In-Memory Data Fabric, which expand the benefits of IMC to the Hadoop world by enabling enterprises to achieve unmatched performance and scale with their existing MapReduce applications. All this is possible without requiring any code change to the native MapReduce, HDFS and YARN environment.

Prior to this offering, running IMC in an existing Hadoop environment required code changes to the application, reducing organization's ability to quickly derive the full performance benefits of an in-memory architecture. The In-Memory Accelerator for Hadoop allows for true plug and play deployment, meaning that within minutes of download, developers can deliver up to 10x performance improvement on their Map/Reduce applications.

By enabling companies to more readily leverage in-memory performance at scale for their Hadoop clusters, GridGain is extending the benefits of its proven IMC platform to a larger enterprise community.

GridGain's Hadoop acceleration is based on dual-mode, high-performance in-memory file system that is 100% compatible with Hadoop HDFS — and an in-memory optimized MapReduce implementation. GridGain's in-memory MapReduce effectively parallelizes the processing of in-memory data stored in GGFS. It eliminates the overhead associated with job tracker and task trackers in a standard Hadoop architecture while providing low-latency, HPC-style distributed processing.

By enabling companies to more readily leverage in-memory performance at scale for their Hadoop clusters, GridGain is extending the benefits of its proven IMC platform to a larger enterprise community.

For more information about GridGain and its products, please visit www.gridgain.com.