Analyst Insight

Aberdeen Group
A Harte-Hanks Company

January, 2012

In-memory Computing: Lifting the Burden of Big Data

Business data is growing at an average of 36% per year, according to Aberdeen's research, and organizations of all sizes, across all countries and industries, have to face the challenge of scaling up their data infrastructure to meet this new pressure. Advances in server hardware and application design have led to a potential solution: in-memory computing. In a December 2011 study on the current state of Big Data, Aberdeen's research showed that organizations that had adopted in-memory computing were not only able to analyze larger amounts of data in less time than their competitors - they were literally orders of magnitude faster.

Two Birds with One Stone: Volume and Velocity

As discussed in Aberdeen's recent report Big Data, Big Moves (August 2011), the main problems of the explosion of business data are commonly referred to as the "Three V's": volume, velocity, and variety. Ever-growing **volumes** of data must be stored and accessed, then captured, processed, analyzed and delivered to knowledge-workers with unprecedented **velocity**, in a broad **variety** of different formats.

A complete data management roadmap must address each of these areas, and no single silver bullet can slay the Big Data beast. However, when we examine the intersection of two of these major challenges - volume and velocity - one technology stands apart in its ability to meet these demands: in-memory computing.

On a macro level, almost all Big Data problems are about time. The greater the data volume and the faster that data streams into the enterprise, the longer it takes for traditional analytics and data management software to turn this data into actionable information. According to Aberdeen's December 2011 research study on Big Data, the most commonly cited pressure driving organizations to invest in better ways to deal with the data explosion was that business users could not get the information they needed fast enough (Figure 1). In-memory computing can make the greatest impact in addressing this problem.

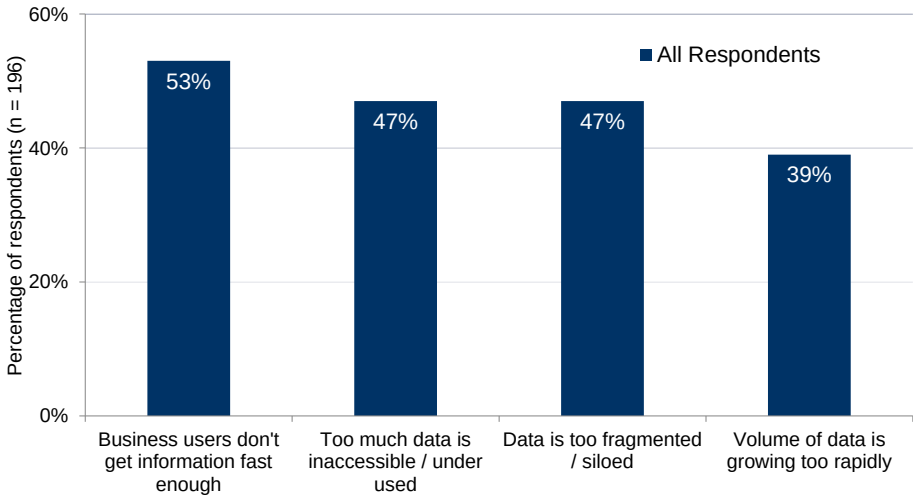
The goal of in-memory computing is to move data as close as possible to processors. Traditionally, when data needs to be analyzed, a data look-up must be performed on the appropriate database, a process that is bounded by the speed of the disk drive it is stored on. The procedure is further slowed by any latency involved in transfering this data through an input / output (I/O) connection between the storage device and server.

Analyst Insight

Aberdeen's Insights provide the analyst perspective of the research as drawn from an aggregated view of the research surveys, interviews, and data analysis

This document is the result of primary research performed by Aberdeen Group. Aberdeen Group's methodologies provide for objective fact-based research and represent the best analysis available at the time of publication. Unless otherwise noted, the entire contents of this publication are copyrighted by Aberdeen Group, Inc. and may not be reproduced, distributed, archived, or transmitted in any form or by any means without prior written consent by Aberdeen Group, Inc.

Figure 1: Pain Points of Big Data



Source: Aberdeen Group, December 2011

In-memory computing, however, relies on the latest breed of high-powered servers with multi-core processors and vast amounts of local random access memory (RAM). Terabytes of critical business information can be stored directly on the server's RAM, allowing for zero latency, no look-up, high speed, multi-core processing. Coupled with applications designed to take full advantage of this leap in available computing power, organizations can now realize the dream of real-time or near real-time analysis of core business data.

Performance Benefits of In-Memory Computing

While the philosophy behind this technology is sound, and the few case studies available have been impressive, in-memory computing is still a fairly new player in the IT field. It can be easy to dismiss the marketing rhetoric around potential performance gains as exaggerations or niche cases. However, Aberdeen's research indicates that the benefits are real.

Aberdeen examined 196 organizations world-wide that are currently dealing with Big Data, of which 33 reported having implemented in-memory computing. These companies represent a large cross-section of industries and sizes, and dealt with business data ranging from a handful of terabytes to multiple petabytes. In aggregate, the organizations that implemented this technology outperformed their peers by two orders of magnitude. They were able to process much more information, much faster and with greater efficiency (Table 1).

Fast Facts

Organizations with in-memory computing reported the following amounts of active business data:

- √ 50% had under 5 TB
- √ 25% had 5 to 250 TB
- √ 12% had 250 TB to 1 PB
- √ 13% had over 1 PB

Organizations without in-memory reported:

- √ 45% had under 5 TB
- √ 34% had 5 to 250 TB
- √ 13% had 250 TB to 1 PB
- √ 8% had over 1 PB

© 2011 Aberdeen Group.

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 Group, December 2011

Due to the high varience in the amount of active business data and the amount of data typically analyzed per organization, Aberdeen used the median value for the two groups. While many organizations in both categories were processing hundreds of terabytes, the majority of organizations dealt with a few dozen. Respondents with in-memory computing show substantially better performance in dealing with the volume of their data. They successfully store larger amounts and analyze more data at once, both in terms of byte size and the percentage of overall business data that could be processed. The real gains, however, came in the velocity category. Organizations with in-memory computing reported an average of 42 seconds for their analytical system to respond to a query - 26% reported a response time under 10 seconds, and 100% reported response times under 5 minutes. This was **over one hundred times faster** than organizations without this technology, with an average response of an hour and fifteen minutes.

However, this technology, despite its strong correlation with incredible performance gains, is not for every organization. The initial investment is steep, requiring the latest in server hardware, along with new software licenses and analytic applications. It comes as no surprise then, that 43% of organizations with in-memory computing reported annual revenue of over $1 billion US, compared to only 28% of those without in-memory computing. Petabyte-scale Big Data problems are mostly an enterprise-level problem, so organizations without in-memory computing are more likely to represent the SMB and SME markets, and the unique Big Data challenges they face. However, the price of entry for in-memory computing will likely decline as economies of scale kick in and production costs decline. Server

Address the quality and consistency of your data at the source or as close to the source as you can. Pay attention to consistent business processes and to naming standards with properly identified global meaning across the enterprise.

~ Manager

$1B - $2.5 Billion North American Chemicals Company

Demographics

Organizations with in-memory computing represented the following company sizes:

√ 43% had > $1 billion annual revenue

√ 17% had $50 million - $1 billion

√ 40% had < $50 million

Organizations without in-memory represented the following:

√ 28% had > $1 billion annual revenue

√ 29% had $50 million - $1 billion

√ 43% had < $50 million

© 2011 Aberdeen Group.

hardware has a long history of rapid price declines, making formerly unaffordable technologies available to SME budgets.

Furthermore, in-memory computing is not designed to address the variety of data formats, such as analyzing unstructured or semi-structured data. Its strength lies in computing large volumes of structured data records, such as transaction data, sales figures, customer information or product codes.

For enterprises wondering if in-memory computing is worth the cost, Aberdeen's research makes a strong case for its value. The average organization using in-memory computing spent over $850,000 US on all their data storage and processing hardware in the last 12 months, and a further $820,000 US on data management and analytical software. However, they were able to use this technology at greater scale, providing 15% of all their employees with access to business data and analytics; organizations without in-memory computing could provide only 9% of their employees with access. Taking the number of employees each group was able to support into account, those with in-memory computing spent 64% less per seat on hardware, and 36% less per seat on software. For organizations needing to process large amounts of data and deliver it to a large number of employees, in-memory computing appears to be a cost-effective option.

Big Data is a huge opportunity to gather insight in what's going on in the enterprises ecosystem, especially customers. We need the technology to analyze and get the results we are asking for.

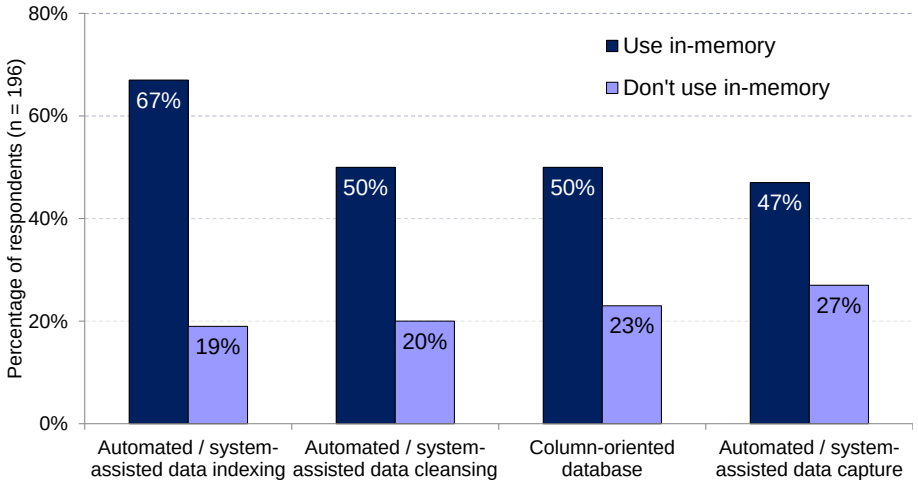
~ IT Manager

>$5 Billion EMEA Software Company

Automation and Scale: The Foundation of Big Data

As mentioned previously, no one technology can solve all of an organization's Big Data woes. In-memory computing is no exception. Organizations only realize the incredible performance reported above with the help of supporting technologies and capabilities.

Figure 2: Technology to Enable Automation and Scale



Source: Aberdeen Group, December 2011

One of the most important aspects of any Big Data intiative is how to address the issue of scale. Business processes that are effective for smaller

© 2011 Aberdeen Group.

companies can quickly become cumbersome bottlenecks when forced to deal with data dozens, hundreds or thousands of times larger and faster. Even relatively simple and quick manual processes rapidly become complex and difficult to manage. For organizations experiencing the growing pains of a rapidly expanding datacenter, or those buried under mountains of under of unused data, automation is often the key.

Organizations that have invested in in-memory computing have mostly adopted this strategy already. They have automated the process of capturing data and ushering it into the appropriate databases or data warehouses. They have a method for classifying and indexing this data for easy use within existing systems and applications, as well as the ability to cleanse and standardize this data so it can immediately be analyzed with older, historical data. Lastly, they are much more likely to have adopted columnar databases. This newer database system is a departure from the older row-based database design and can improve analytical speeds by reducing seek times, and the amount of data that needs to be accessed at once. While not the best fit for every database, this approach has a particular synergy with in-memory computing hardware that analyzes large amounts of data in aggregate. Organizations with in-memory computing are two- to three-times as likely to have adopted each of these technologies as those without.

Technology is the key enabler of a Big Data initiative. Data standards must be defined. Analytics must be powerful and usable by business users directly, rather than depending on IT.

~ Vice President of IT

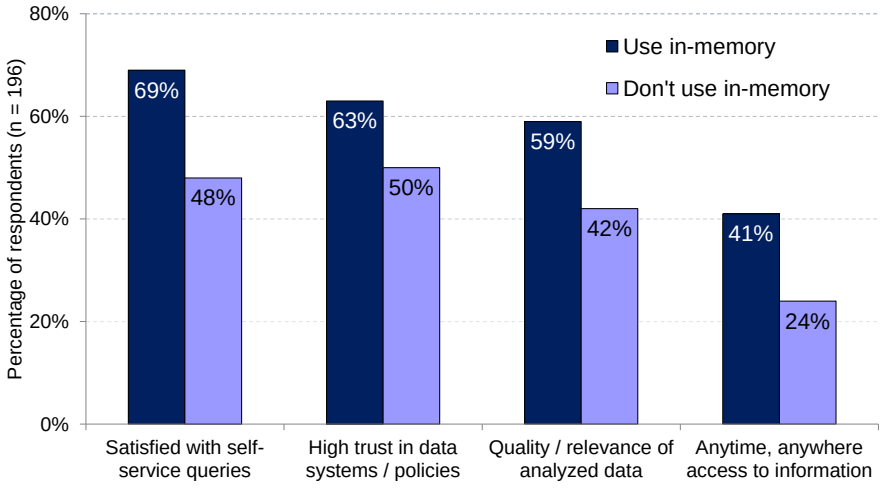
$1B - $2.5 Billion APAC

IT Services Company

Building Trust in Business Data

If the datacenter and IT department represent one side of the Big Data picture, the other is the line-of-business worker. All the speed and performance benefits in the world mean nothing if the data system is too hard to use, or if the information it delivers isn't trusted.

Figure 3: Satisfaction and Trust in Business Data



Source: Aberdeen Group, December 2011

I think the biggest issue around having Big Data bring value is the ability to trust in its accuracy. That is the biggest hurdle to get over. Once that is done, you can get on to the fun part of making it tell a story that is usable for the business owners.

~ Procurement Manager

>$5 Billion North American Pharmacutical Manufacturer

© 2011 Aberdeen Group.

In Figure 3, Aberdeen shows the impact these technology investments had on the knowledge-workers dealing with information every day, and the decision-makers relying on analysis to set corporate direction. In the December 2011 research study on Big Data, organizations were asked to rate the average knowledge-worker's opinion of various data-centric business processes on a scale of 1(very unsatisfied) to 5 (very satisfied), with a 4+ receiving a rating of "satisfied". According to Aberdeen's research, organizations with in-memory computing had employees with higher levels of trust in both the data systems and resulting information

Trust and satisfaction are not only important for morale, but they also have a large impact on data management policies and procedures. When employees believe (correctly or not) that information can't be trusted, or that a system is too frustrating or unreliable, they often turn to work-arounds. As mentioned in Data Quality and the Supply Chain (October 2011), these non-sanctioned manual efforts can not only take longer, but can put data at risk for security or non-compliance incidents, and can introduce more human error into business processes.

Finally, organizations with in-memory computing reported higher levels of satisfaction with self-service data queries. While less than half of these organizations (41%) also reported high satisfaction with anytime, anywhere access to business information, this was almost twice the rate of other organizations. Both these elements point to a growing trend in business: the desire for business information to be accessed anywhere, shared instantly, and managed from a central, trusted location. Aberdeen will be exploring this concept of the convergence of Social, Mobile and Cloud technologies, or SoMoClo™, over the coming months. According to this latest research, it appears that organizations with in-memory computing are more likely to be on the forefront on this movement.

Fast Facts

- √ 41% of organizations with in-memory computing had Mobile BI tools vs. only 11% of other organizations
- √ 59% of organizations with in-memory computing had a role for evaluating new data management technologies. Only 38% of other organizations also did so.
- √ 59% of organizations with in-memory computing had a dedicated data scientist / analyst vs. 44% of other organizations

Summary and Recommendations

In-memory computing is particularly suited to alleviate two of the major Big Data pain points many organizations are experiencing: volume and velocity. When it comes to processing large amounts of structured data very quickly, organizations with in-memory computing were reported to vastly outperform their peers. They were able to analyze over three times the amount of data at over a hundred times the speed of other organizations.

By tackling the volume and velocity aspects of Big Data, employees at organizations with in-memory computing reported higher levels of satisfaction and trust in the data systems and resulting analysis. As Aberdeen's research has shown time and again, higher quality data correlates directly to improved business processes.

In-memory computing is most applicable for large enterprises, and can have a hefty price tag for the initial investment. However, with data volumes continuing to grow by an average of 36% per year, organizations must take steps to master that information. In-memory computing provides the muscle to get the job done.

© 2011 Aberdeen Group.

Aberdeen Group
A Harte-Hanks Company

For more information on this or other research topics, please visit www.aberdeen.com

Related Research	
How Much of Your Data Should be in the Public Cloud?; November 2011 *Data Quality and the Supply Chain: The Impact of MDM and Portals*; October 2011 *Big Data, Big Moves*; August 2011 *Archive Agility: Building Business Resilience through Active Archiving*; July 2011	*Turning Pain into Productivity with Master Data Management*; February 2011 *Data Management for BI: Fueling the Analytical Engine with High-Octane Information*; December 2010 *Predictive Analytics: The Right Tool for Tough Times*; February 2010

Author: Nathaniel Rowe, Senior Research Associate, Information Lifecycle Management; (nathaniel.rowe@aberdeen.com)

For more than two decades, Aberdeen's research has been helping corporations worldwide become Best-in-Class. Having benchmarked the performance of more than 644,000 companies, Aberdeen is uniquely positioned to provide organizations with the facts that matter — the facts that enable companies to get ahead and drive results. That's why our research is relied on by more than 2.5 million readers in over 40 countries, 90% of the Fortune 1,000, and 93% of the Technology 500.

As a Harte-Hanks Company, Aberdeen's research provides insight and analysis to the Harte-Hanks community of local, regional, national and international marketing executives. Combined, we help our customers leverage the power of insight to deliver innovative multichannel marketing programs that drive business-changing results. For additional information, visit Aberdeen http://www.aberdeen.com or call (617) 854-5200, or to learn more about Harte-Hanks, call (800) 456-9748 or go to http://www.harte-hanks.com.

This document is the result of primary research performed by Aberdeen Group. Aberdeen Group's methodologies provide for objective fact-based research and represent the best analysis available at the time of publication. Unless otherwise noted, the entire contents of this publication are copyrighted by Aberdeen Group, Inc. and may not be reproduced, distributed, archived, or transmitted in any form or by any means without prior written consent by Aberdeen Group, Inc. (2011a)