OVUM

Accelerating time-to-insight for midsize companies using in-memory analytics

Written by: Helena Schwenk   Published April 2010, © Ovum

CONTENTS

EXECUTIVE SUMMARY

OVUM VIEW

In-memory analytics provides business users with a fast and streamlined way to access and analyze information. The ability to store and process data in-memory facilitates on-the fly analysis and reporting, rapid querying and the slicing and dicing of large data sets. Because of this, users can run queries and generate reports in seconds that would otherwise take minutes or even hours with conventional tools. The faster response times that in-memory analytics affords means businesses can become more agile, make quicker decisions and accelerate time to insight.

These are important benefits for midsize companies; many of whom face the same business issues and challenges as large enterprises but lack the IT skills and resources to tackle them using traditional big iron and expensive BI and data warehousing solutions. Moreover, improving time-to-insight is becoming increasing important for midsize firms as changes in the speed and pace of business and the impact these have on the requirements for storing, analyzing and acting on high volumes of information is set to increase.

Importantly in-memory technology can also reduce reliance on IT and in so doing improve end user self serviceability. This is because the tasks of defining, building and pre-calculating an On-Line Analytical Processing (OLAP) cube are greatly simplified. OLAP analysis provides many benefits to midsize companies as it allows business users to interactively explore and identify trends from important business measures (such as sales, profits and costs) along many different dimensions (such as products, channels and time). Improved self-service is particularly beneficial for midsize companies for whom pressure on IT resources and budgets and/or a lack of technical expertise can cause significant challenges in meeting business goals and objectives. With an in-memory approach there is less need to define upfront the types of business questions and/or queries that need answering through predefined indexes, calculations, query drill paths or summarized data – all of which are typical activities in a conventional OLAP environment.

The viability of in-memory analytics is also closely linked to advances in hardware technology, such as 64-bit computing, multicore and improvements in processor speed. These technical advancements are helping companies optimize the use of memory and speed up data processing performance over previous incarnations of the technology. The combination of more scalable, faster in-memory technology coupled with falling memory prices significantly improves the price/performance ratio of in-memory analytic solutions, helping to propel it as a viable cost effective technology for midsize organizations.

In-memory analytics also has the potential to open up analytics to more users, by allowing them to ask more pertinent questions out of curiosity, gain greater insight, derive actionable information and do this in increasingly faster timeframes which can only be beneficial for the business. The bottom line is that the faster implementation timescales afforded by shorter and less costly development cycles makes in-memory analytics a very cost effective tool for the decision support needs of midsize companies today.

BUSINESS ANALYTIC REQUIREMENTS OF MIDSIZE COMPANIES

Midsize companies are characterized by their unique challenges, business focus and processes; traits that also impact and influence the downstream requirements of any analytics project undertaken. While midsize firms have many things in common with large enterprises they radically differ in terms of the level of budget and resources they have at their disposal. Not surprisingly this is a major factor governing the choice and selection of any BI reporting, analysis or planning solution.

When a midsize company is selecting a business analytic solution, IT departments are looking for technology that is cost-effective and faster to deploy. Hence Total Cost of Ownership (TCO), speed of deployment, and reducing risk all factor into that all-important Return On Investment (ROI) equation. Additionally, the ability to provide pre-packaged content based around a company's core industry focus and business processes – one that require less technical expertise to set up and use – is also likely to provide high appeal to midsize firms that do not have deep analytic modeling and analysis skills in-house. This is part of the reason why technologies and deployment models such as in-memory, pre-configured appliances and software-as-a-service (SaaS), that require less upfront capital investment and provide more predictable IT costs, are gaining implementation momentum in midsize firms.

EVOLUTION OF IN-MEMORY ANALYTICS

PROVIDING FASTER WAYS TO ACCESS INFORMATION

The term 'in-memory' describes the notion of loading and processing data in Random Access Memory (RAM) rather than on disk. One of the key technical features that differentiates in-memory from conventional on-disk storage methods is its lower I/O burden and reduced CPU cycles. This makes accessing data within memory an order of magnitude faster than accessing it within an on-disk database system. The design philosophy of in-memory approach brings many benefits to business analytic applications, especially those that demand fast data access, storage and analysis – a capability that, to a large degree, would be cost-prohibitive today for many midsize companies in terms of software, hardware and services investments.

HARDWARE ADVANCES ARE MAKING IN-MEMORY MORE VIABLE

In-memory technology has been around for a couple of decades, vendors such as Applix with its TM1 product (which was acquired by IBM Cognos) was one of the early pioneers of in-memory processing. Notwithstanding its lineage the core value proposition of the technology still works now as it did then – i.e. the ability to access, process and analyze data for faster and speedier insight. However, the difference today is that in-memory technology is becoming more scalable, faster and cost effective than its earlier counterparts. One of the key factors influencing this change is the declining cost of RAM. But equally the popularity of in-memory can also be linked to incremental hardware advances that are helping to lower barriers to adoption and improve the price/performance ratio of the technology. These include improvements in memory capacity, processor speed, multi-core and massive parallelism which optimize the use of memory, eliminate previous database size limits and speed up data processing performance over previous incarnations of the technology.

The combination of the declining cost of RAM with these technological innovations is helping in-memory move out of its current niche status – within certain vertical and application areas – into a more viable and cost effective mainstream technology platform. Likewise, as vendors continue to harness these hardware capabilities, in-memory technology is gradually evolving to represent a more scalable, low-latency and arguably more cost effective technology for midsize companies to use. As a consequence an increasing number of BI vendors are starting to incorporate in-memory designs into their platforms - either as core components or as optional modules.

ADVANTAGES OF IN-MEMORY ANALYTICS

FASTER SPEED OF RESPONSE

The pervasive use of the internet continues to raise the bar for levels of information accessibility and responsiveness. These increased expectations also pervade other IT systems (including business analytic solutions), with many business users expecting equally easy, fast and reliable access to business information. However the reality today is that many BI users still contend with slow response rates or query performance. Often they are compelled to seek other means to access the information they need such as using the ubiquitous Excel. Unsurprisingly this often translates to users extracting, loading and reporting off their own data sets and information silos within bespoke PC databases or spreadsheets. While this might satisfy a narrow, local analytic need, it carries the danger of perpetuating data management issues associated with the creation of separate islands of information and multiple versions of the truth, which can easily undermine an company-wide business analytics implementation.

The analytic needs of midsize companies are typically no less complex than those of large enterprises; although the data volumes might be less voluminous. So it's not surprising that a midsize firm equally recognize the need for speed in analytics. It is also no coincidence that fast data query and analysis is cited more than any other feature as being most important for business analytic software buyers. Yet conventional BI tools – typically designed to function around batch-oriented processing models – can only go so far, and often fail to keep pace with decisions that have to be made in the spur of the moment. Frequently they lack capabilities that allow users to interactively work with data and follow a train of thought; instead users are left waiting sometimes minutes, hours or overnight for queries to run.

This requirement for fast, rapid access to reports, analysis and business metrics at near real-time speeds can be satisfied through in-memory analytics. The technology facilitates fast querying, slicing and dicing of large data sets and on-the-fly calculations without the need to resort to more traditional methods such as summarizing data, pre-built cubes, IT-intensive database tuning or the laborious and constant task of fine tuning of queries. That not only means faster analysis – users can perform adhoc queries in seconds that would take minutes or even hours with conventional tools – but also faster self-service BI without users having to wait in the IT queue.

IMPROVING SELF SERVICE THROUGH ANALYTIC FLEXIBILITY

Business managers want (and expect) to analyze and report on the state of their business today rather than rely on someone else to do it. While some midsize companies are still a far cry away from realizing that vision, advances in-memory technology and a continued rise in IT literacy among savvy business users are helping many move closer to achieving this goal. Reducing reliance on IT through improved self serviceability can only be beneficial to midsize companies for whom pressure on IT resources and budgets and/or a lack of technical expertise pose significant challenges to IT deployment.

The lightning-fast query response times and ability to access and analyze data on-the-fly makes in-memory analytics particularly suitable for operational business reporting and monitoring. By loading data in-memory from operational data sources, front line users can also be empowered with increasingly timely and drillable information about the state of business now, as opposed to yesterday or last week, so they can for example monitor sales trends, identify problems in the supply chain or pinpoint under performing products. It is the immediacy of analytic insight and rapid BI response that is helping companies increase the speed to insight and lower the latency barriers for BI.

In-memory technology can also fast track analytic processes as the increased speed of query response allows users to reformulate new queries and explore new lines of thinking without having to stop and pause for a query to return its results. Some in-memory analytic tools are not constrained by an underlying pre-built data model allowing users to perform instantaneous exploratory analysis and ask questions that may not have previously been possible. In addition, the ability to source raw data and visually define a cube for on-the-fly analysis is a valuable tool for business users especially if the modeling tool is intuitive enough.

SUPPORTING SPECIALIZED BUSINESS ANALYTIC REQUIREMENTS

The fast underlying architecture of in-memory also provides a useful platform for what-if analysis, enabling users to model and simulate different scenarios and assess the impact on results. This is advantageous in today's fast moving and dynamic economic climate as it helps business users represent business situations quickly and ensures they are not caught out by sudden changes in market direction. Likewise within a financial planning context, the ability change plans frequently to analyze the potential impact of changing market conditions and opportunities such as change in interest rates or increase in wholesale prices is a particularly valuable asset. For example, by drawing up best-case and worst-case scenarios, it becomes possible for business users to gain an understanding of the implications of each scenario and makes it easier to understand the different choices that may have to be made, and the range of outcomes possible.

While what-if analysis isn't particularly unique in planning cycles, the ability to combine it with real time write-back - by leveraging in-memory technology - is. While Excel and on-disk OLAP applications can support write-back (for example by building macros or recalculating values within the OLAP cube) they do not provide the same level of flexibility and rapid response rates of an in-memory environment. Real time write back can be an extremely useful feature in planning, budgeting and forecasting applications not least through its ability to make financial information more realistic, relevant and timely but because it can help bridge the gap between insight and action.

Just as spreadsheets enable users to write new values back into a cell and examine how this value affects other cells, similarly in a more complex multidimensional planning model users may wish to write back values in real time to explore the consequences. For example a finance user may want to change values in a budget to look at the real time impact on revenues, expense and cash of a potential acquisition or the sale of a subsidiary or division. This combination of both in-memory technology together with write-back functionality can be a very useful feature for analysts as it provides both interactivity and immediacy of results.

REDUCING THE IT BURDEN

A key benefit of in-memory analytics is it eliminates the need for a pre-built OLAP cube (MOLAP) or relational data mart (ROLAP) as the whole cube or datamart is ideally loaded into memory. This removes a lot of the cost, skill and resource requirements associated with developing and building a conventional data warehouse. Instead with in-memory the tasks of defining, building and pre-calculating an OLAP cube are greatly simplified as there is no need to define upfront the types of business questions and queries that need answering through predefined indexes, calculations, query drill paths or summarized data for example. In-memory simply allows companies to integrate data from transactional systems, external data sources, spreadsheets or existing relational data warehouses and model and load it into memory ready for querying and analysis. As a result, in-memory can dramatically accelerate project implementation timescales, improve the time to value and lessen the amount of data modeling and data preparation needed. However it is important to note that these benefits can only be truly realized if the in-memory model works against a clean and consistent set of data. If it doesn't then it simply helps users make the wrong decision quicker.

This is in contrast to traditional analytic approaches that require developers to pre-build a multi dimensional OLAP cube or datamart and define upfront the types of business questions and queries that need answering through pre tuned schema design, indexes and pre built calculations. For conventional data mart environments that rely on pre defined database indexes to accelerate queries and response times this can also significantly increase the physical size (and cost) of the data mart as the indexes are often larger than the raw data itself. This is something that can be averted with an in-memory approach since there is less need to build performance boosting indexes as calculation and aggregates can be built on demand. Furthermore, many vendors employ sophisticated data compression techniques that substantially reduce the data footprint of in-memory solution in comparison with conventional on-disk methods.

When the requirement to build and maintain a multidimensional cube or datamart is removed, IT spends less time and effort on setting up and supporting the overall analytic system. This gives IT more time to focus on the other urgent development and maintenance tasks vying for their attention. For over stretched IT departments this could also mean freeing up highly sought after DBA skills or helping reduce the specialist resource requirements of a BI and analytics system.

DEPLOYMENT CONSIDERATIONS

Deploying business analytics technology carries both business and technical challenges and some risk. Therefore, organizations considering an in-memory analytic deployment need to be equally carefully by planning out implementations and understanding the limitations of its use.

ENSURING A COMPLEMENTARY APPROACH TO EXISTING BI DEPLOYMENTS

A significant amount of in-memory analytics solutions have been deployed on a standalone departmental or data mart basis either as a result of their shorter implementation timescales or their particular business focus. However, improvements in data and user scalability, metadata management and end user functionality mean in-memory technology and tools are now well positioned to grow and evolve in line with changing business requirements. While (on-disk) data warehouses still remain a viable option for midsize firms who want to transform, cleanse and consolidate information from across the enterprise, an in-memory solution can also complement an existing data warehouse. A common scenario, for example, is when users complain about query and report performance or that it takes too long for IT to respond to their information requests. This is when an in-memory solution could act as a high performance, targeted down stream BI engine designed to provide a subset of data for a particular use case or groups of users. Hence users benefit by getting access to a faster analytic environment and IT can service the information needs of a subgroup of users without the need to build a separate cube or data mart, summarize data or get involved with time consuming data preparation tasks. For midsize companies who are just getting started on their business analytics journey then in-memory could provide a great starting point as it allows companies to get up-and-running with less effort, time and cost than a larger data warehousing implementation.

However, a word of caution. Those organizations considering deploying an in-memory database separate from their data warehouse (usually as a result of wanting to speed up deployment times) must be careful not to create yet another 'version' of the truth with its own separate business logic, definitions and calculations. Organizations must therefore look to strike a balance between both approaches through appropriate tool support at the data integration, metadata and query layers. The trick is to ensure that both architectures can co-exist without compromising the business value and integrity of the other. Ovum recommends that midsize companies seek advice and guidance from their vendor of choice for the availability of best practice or case studies that illustrate how both environments can co-exist successfully.

GUARANTEEING DATA QUALITY WITHIN IN-MEMORY ENVIRONMENTS

All business analytics solutions need a solid, robust and trustworthy data foundation to stand any chance of successful adoption. The best designed analytics system will fail if it is not underpinned by a strong data management foundation where the data is accurate, fresh and integrated from multiple sources, with consistent definitions that are agreed upon. An in-memory analytics system is no different.

Indeed it raises the stakes for data quality. If data is loaded directly from operational source systems into memory careful consideration should not only be given to controlling how data is moved, integrated, accessed and used. More important perhaps is the quality of the data sources being accessed, in terms of cleanliness (hygiene), referential integrity, consistency, accuracy etc. If data is sourced from a data warehouse then the burden of cleansing the data is somewhat alleviated and only guaranteed by the underlying data quality processes in place. Whatever the data acquisition process the end goal of any analytics solution is to deliver trustworthy and usable business information to in order to generate actionable business insights.

IN-MEMORY ARCHITECTURAL APPROACHES VARY

Although in-memory tools share many feature-specific traits, they vary considerably in terms of their architectural approaches. Understanding these different approaches can help organizations make an informed choice about the suitability of the technology to their particular environment and business requirements.

The different approaches to in-memory are outlined in table one.

Table 1: The different approaches to in-memory	
Associative model	Loads and stores all of its data in an 'associate' data model that runs in-memory. Associations are based the relationships between various data elements so when a user to clicks on an item within a data set, the selected item then turns green and all associated values turn white. This allows users to quickly query all relevant data without the dependency of a predefined hierarchy or query path and are not limited to navigating the analytical data in a preordained way
In-memory OLAP	Works by loading data in-memory allowing complex calculations and queries to be computed on-demand resulting in fast response times. If write-back is supported then users can change assumptions on the fly to support real time 'what-if' scenarios a specific requirement in financial planning and forecasting.
Excel in-memory add-in	Allows users to load large volumes of data into Excel using in-memory. Once the data is within Excel relationships are automatically inferred between the data sets allowing users to perform on-the fly sorting, filtering and slicing and dicing of large data sets which overcomes some of the technical data volume limits of Excel. This approach is designed to improve self-service capabilities as it reduces reliance on IT and lessens the need for business users to become proficient in multi-dimensional structures and techniques. The Excel in-memory add-in is dependant on a particular back-end data management and portal platform which aids data sharing and collaboration.
In-memory accelerator	Designed to improve query performance within an existing BI environment. The accelerator works by loading data into memory and leveraging pre built indexes to support super fast query response times. Since this approach does not tie the analysis down to any predefined paths or hierarchies, users have greater analytic latitude and freedom to explore the data. Some in-memory accelerators are also dependent on the use of highly parallel hardware which will need configuring and optimizing before use.
In-memory visual analytics	Combines an in-memory database with a visual data exploration tool allowing users to quickly query data and report results within a visual and interactive analytics environment.
Source: Ovum	**O V U M**

DATA MODELING AND METADATA FLEXIBILITY IS KEY

How data is modeled within an in-memory environment, and the level of complexity involved, is a key concern for the IT department. Designers of the in-memory business model need tools that offer enough flexibility for the model to be built to fit the business needs. Part of the process of defining a business model involves creating the associated dimensions, cubes, and business rules. Sometimes this is on a per-model basis but a better alternative is to create a metadata business model that provides an abstracted common understanding and view of the data. The advantages of the second approach are that it ensures consistency of terms, especially across hierarchies, business rules and calculations across the multiples part of the in-memory solution. This abstraction layer brings many benefits to both business and IT users. The most important ones being, it reduces duplication of effort so users avoid recreating metadata or new data silos with every application, it is easier to manage because it decreases the downstream effects of changes and it enables power business users to perform ad hoc analysis against any data set. Ideally this final point would be happen without the need for scripting.

CHECK VENDORS 64-BIT SUPPORT

The adoption of 64-bit computing has increased the amount of data that could be analyzed at any one time with in-memory technology, eliminating the previous 4GB addressable memory space limit of 32-bit operating systems. While 64-bit operating systems are not necessarily a prerequisite for an in-memory environment they provide substantial performance and scalability gains over the 32-bit version. However organizations should be always check the level of vendor 64-bit platform support as older hardware and certain operating systems may not be supported with all 64 bit systems.

SUMMARY

Midsize firms have many things in common with large enterprises however they differ radically in terms of the level of budget and resources they have at their disposal. Not surprisingly this is a major factor governing the choice and selection of any business analytics or planning solution. In-memory analytics provides midsize companies with a faster, more flexible, and arguably lower-cost way of accessing and processing information allowing users to get answers to business questions in seconds rather than hours. By virtue of its high performance architecture in-memory has the potential to help midsize organizations become more informed, agile and respond quicker to changing market conditions.

In addition, advances in technology and lower costs of memory and CPU make this type of technology more attractive than ever before. Matching the appropriate architectural approach with the kind of business analytics solutions needed by a midsize company has the potential to deliver benefits such as reduced time to insight, greater agility, increased self-service and lower overall IT demands.