Why an integrated platform is the key to real-time insights

Real-time insights have become a necessity for today's demanding business leaders as they look to increase operational efficiency, improve profitability, reduce costs and gain a competitive edge in the marketplace.

Find out more about *Performance* at **ey.com/performance.**

The drive for real-time insights is relevant for organizations of all sizes and across all industry segments.

The types of sector-specific issues are far-reaching but all have one thing in common: the need for real (or near real) time business insights. For example, manufacturing companies are looking for effective ways to estimate warranty claims, enabling them to be accurate in setting warranty reserves. Media and entertainment companies are seeking a means to predict box office success and help them make informed decisions on film marketing, promotion, distribution and investment.

In the financial services sector, firms have a list of wants, such as a better avenue to determine if a specific customer poses a potential credit

risk, the means to enrich and further refine customer segmentation, and the need to improve profitability and reduce reserve requirements through more accurate financial forecasting.

Retailers require information to better predict the demand for specific products at different times of the year based on factors such as weather or seasonality. And telecommunications firms are looking to gain insight into customer churn and payment rates.

All of these insights can only be obtained by amassing, storing and analyzing large volumes of both structured and unstructured data.

EY
Building a better working world

Drowning in data, paralyzed by technology

While many leading organizations are drowning in data, they lack the ability to analyze and drive value from the vast amount of information at their disposal. Furthermore, many organizations are unable to drive actionable value from their legacy technology investments and recent technology acquisitions. The rapid influx of technology appears to have (in some cases) paralyzed organizations when it comes to understanding how these solutions can coexist within their organization and help them harness the available data both inside and outside their firewall.

Having an end-to-end, integrated big data and analytics platform is critical to an organization's ability to gain real-time (often game-changing) business insights – and to do so in a cost-effective manner.

Simplifying the complex

So what is an end-to-end, integrated big data and analytics platform and why is it so important?

End-to-end refers to the entire life cycle of data, from identification to insight. The key is in spotting and understanding what data is relevant to solve a specific business issue. Data acquisition tools and processes are then required to integrate and store this data so that it can be validated, cleansed and, finally, analyzed to provide true insights.

Integration is the process of acquiring and delivering data to the right data management platform at the right time so that analytics can occur at the business' own pace.

For example, leveraging technology such as Hadoop[1] to store massive quantities of data (structured, unstructured and semi-structured) offers an organization the opportunity to reduce storage costs and the ability to acquire complex and diverse data (social, sensor, cell center, etc.) via batch, real-time and streaming inputs.

Initially, Hadoop was adopted by organizations to solve IT-centric issues, such as the storage of large data sets of varying data types. The more innovative solutions around data discovery and analytics came from leading emerging technology companies (rather than the simple, open-source Hadoop). However, the focus of leveraging Hadoop has dramatically shifted to driving enterprise business insights.

The challenge is that open source solutions such as Hadoop can be somewhat cumbersome to administer and require a fairly unique set of programming skills. Additionally, there is something of an application gap in the Hadoop ecosystem creating some frustration with the pace of development of Hadoop-based applications. Where the real value is created for business application and use, therefore, is in solutions that provide the advanced analytics capability to demystify the MapReduce[2] layer.

Finding the right blend

While Hadoop offers a cost-effective means to address data volume and variety, it tends to fall a bit short in terms of the third "V": velocity (the ability to easily analyze data in real-time).[3] This is where in-memory solutions (available from leading software providers such as Oracle and SAP) can add tremendous value.

SAP HANA, for example, is a true in-memory database solution where data indexes and other artifacts run entirely in the physical memory. This design allows for real-time analytics across large volumes of data. Furthermore, in-memory solutions such as SAP HANA can and are being used as the technology foundation

1. Apache Hadoop is "open-source" (i.e., available at no cost for use or modification by the general public) software for storing and large-scale processing of data, ibm.com/software/data/infosphere/hadoop/, accessed November 2014.

2. The term MapReduce refers to two separate and distinct tasks that Hadoop programs perform. The first is the map job, which takes a set of data and converts it into another set of data, where individual elements are broken down into tuples (key or value pairs). The reduce job takes the output from a map as input and combines those data tuples into a smaller set of tuples. As the sequence of the name MapReduce implies, the reduce job is always performed after the map job. ibm.com/software/data/infosphere/hadoop/mapreduce/, accessed November 2014.

3. The "3Vs" is an accepted model used for describing big data, credited to D. Laney, "3-D data management: controlling data volume, velocity and variety," Gartner, 2001, blogs.gartner.com/doug-laney/deja-vvvue-others-claiming-gartners-volume-velocity-variety-construct-for-big-data/, accessed November 2014.

to extend business applications (i.e., SAP Business Suite on HANA) where "hot" data is required and needed by the business. However, in-memory solutions (can) come up short in terms of cost-effectively storing massive quantities of data and do not handle complex, unstructured data as well as Hadoop-based solutions.

The end-to-end, integrated approach to data and analytics is a blend of the low cost of data acquisition and storage of Hadoop coupled with the unprecedented speed and agility of in-memory platforms.

Unlocking value
The opportunity for businesses, therefore, is that they are now able to manage data in a way that reduces the overall cost associated with a traditional data warehouse while aggressively exploiting new analytical capabilities. This enables management to turn poor business decisions, made using haphazard guesswork, into well thought out and successful business decisions that lead to execution and actions.

An end-to-end, integrated big data and analytics platform unlocks the value of today's leading emerging technology solutions, such as in-memory and Hadoop, and paves the way for advanced analytics (analytics that provide for the forecasting of future events and behaviors, allowing businesses to perform what-if analysis to predict the effects of potential changes in business strategies).

When coupled with core and legacy business applications (ERP, CRM, HR, etc.), and leveraging best practices and processes for the effective use of each in a collaborative way (really relying on robust data governance), the integrated business intelligence environment and platform can offer key business advantages and insights never before possible.

For example, in a highly regulated sector such as energy, a company needs to manage pipeline integrity. The ability to leverage predictive asset maintenance (PAM) to foresee when there might be a catastrophic asset failure, such as a broken valve, fitting or pipe, would require the acquisition and analysis of large volumes of unstructured data sourced from sensor devices. In addition, structured data such as manufacturer specifications, as well as historical information on similar types of assets leveraged across the organization, deliver a composite data portrait that anticipates failure. The volume of (diverse) data would be massive (potentially petabytes).

To minimize cost but also enhance the ability for real-time analytics on this invaluable data, an organization would want to leverage Hadoop and an in-memory solution such as SAP HANA. When coupled with data integration tools (potentially tied to back-end systems – such as a supply chain management solution to automate the part replacement process) and business

intelligence reporting and advanced analytic solutions, the result is a value-driven, end-to-end, integrated big data and analytics platform.

An intelligent approach
Deploying value through analytics must be intentional: aligning business with IT and data with its ability to be efficiently acquired, stored and used to drive analytics. Traditional business intelligence platforms alone can no longer keep pace with the needs that shape changing commerce. Emerging technologies, such as Hadoop, are a prerequisite to lower-cost data management. In-memory solutions, such as SAP HANA, offer speed-of-thought analytics. When combined and integrated across the entire organizational data platform, true, value-added, integrated, end-to-end business intelligence is achieved.

Author
Scott Schlesinger
IT Advisory
Partner/Principal, EY, US