The Rise of the Analytic Platform in Big Data

An ENTERPRISE MANAGEMENT ASSOCIATES® (EMA™) White Paper
Prepared for ParAccel
March 2012
# Table of Contents

Executive Summary .................................................................................................................. 1  
The Evolving Enterprise Data Landscape .................................................................................. 1  
Opportunities in Big Data .......................................................................................................... 1  
Managing Hybrid Environments for Performance ...................................................................... 2  
Leveraging Big Data, Analytic Platforms and Traditional Systems ........................................... 3  
  Supply Chain Management in Consumer Retail ........................................................................ 3  
  Financial Services CRM ......................................................................................................... 4  
Strategies for Success ................................................................................................................ 5
The Rise of the Analytic Platform in Big Data

Executive Summary
Analytic driven companies are entering an exciting period of opportunity in data management. Traditional environments are expanding to incorporate new data sources and performance oriented platforms that can process complex workloads and advanced analytics at a scale previously prohibited by both technology and economics. Big Data is delivering a new and deeper view of information at petabyte scale and is relying on surrounding technology to fully leverage and deliver this insight to power the enterprise. Analytic platforms have found a strategic position between the traditional foundations of the Enterprise Data Warehouse (EDW) and Big Data to bring these insights to the optimal platform for delivery and value.

The Evolving Enterprise Data Landscape
Data ecosystems are growing well beyond the confines of the standard EDW environment. New complex analytic workloads and massive data growth are colliding to extend the environment beyond traditional limits in order to meet the analytic needs of today’s sophisticated end users. These new environments include cloud data and its applications, analytic platforms, Big Data platforms and frameworks and even highly available archival systems. Companies are also introducing new data types to add value to their analytics and are seeking new solutions that will allow them to benefit from this information for faster and more insightful decisions.

In short, our historically well-defined data landscape is undergoing a massive overhaul as new data and workloads find the best possible platforms for success. This hybrid environment creates new data management challenges, but foremost it delivers the flexibility, opportunity and value that has been lacking in a more traditional approach to business intelligence and analytics.

Opportunities in Big Data
Big Data is a popular example of data and workload finding an optimal platform. Often defined as data sets that can no longer be easily managed or analyzed with traditional or common data management tools, methods and infrastructures, Big Data carries the characteristics of high velocity, high volume and in some cases a wide variety of data structures that can add to the complexity of managing or utilizing the information. Generally Big Data is delivered via NoSQL or New SQL databases or frameworks. There are hundreds of these platforms to choose from in the market and each carries different performance characteristics and features. For purposes of this paper we will focus on Hadoop – perhaps the most popular of the group.

Many companies are adopting NoSQL and New SQL platforms as part of their analytic strategies and have begun to integrate massive amounts of data into these environments to deliver deeper insights for their analytic consumers. It should be stated that while Big Data is especially well suited for these platforms, extracting value from the process is often done in partnership with analytic platforms or relational databases.

Supporting a Big Data environment for analytics is a group effort between all three data management environments (NoSQL, Analytical and Traditional EDW).
The Rise of the Analytic Platform in Big Data

Companies who have had the forethought to include both NoSQL and Analytic Platforms in their repertoire can reap rewards by managing exactly how best to store, process and execute analytics on the data by leveraging one or more of these solutions in unison.

Managing Hybrid Environments for Performance

Analytic workloads are becoming more complex; in order to achieve optimal output companies need to determine which platform is best for each data type and workload. Often the workload can be shared or divided between platforms to enable the highest level of performance. This is especially true for Big Data analytic projects where the result needs to be combined with additional enterprise data to create value.

When comparing the capabilities of Hadoop Environments, Traditional EDWs and Analytic Platforms you need to consider which is best suited for the storage, processing and final analysis for each project because each platform represents a core competency and value proposition to the overall process. Final analysis is often best completed by analytic platforms that can integrate the data from the other environments and execute additional complex analysis to provide final insights to the user.

This practice provides the enterprise with the best value and best placement of workload and data. Advanced query functions available from some analytic platforms can access results from within Hadoop or EDW systems without physically moving the information. These analytic platforms push queries to these other systems and utilize the results within the analytic process. This ensures that the analysis is leveraging the most current data and eliminates unnecessary ETL complexity.
Leveraging Big Data, Analytic Platforms and Traditional Systems

Retail and financial services are two industries that have embraced the opportunity value of Big Data. Each accesses massive sources that can provide a unique view of the business when coupled with additional enterprise information. The following business examples bring into relief how the analytic platform is playing a highly valuable and strategic role in delivering Big Data value.

Supply Chain Management in Consumer Retail

Consumer Retail organizations often focus on streamlining inventory distribution when seeking competitive advantage. This optimization process needs the ability to match goods with customer demand at the lowest cost to deliver profits via both in-store inventory levels and direct-to-consumer channels.

To address this challenge retailers link operational systems and NoSQL platforms to analytical platforms to execute inventory and demand projection analysis that combine historical performance data, available product levels and consumer sentiment to make highly optimized process decisions.

To gain insight retailers tap into consumer sentiment information available from social media outlets such as Twitter, Facebook and blogs. This sentiment “signal” can then be integrated with supply chain data plus point-of-sale information to make highly informed inventory decisions.

Sentiment data is often highly unstructured making it well suited for Big Data Hadoop environments and less ideal for processing in a traditional EDW or analytical platform. Sentiment data can then be correlated with existing customer transactions from point-of-sale systems and historical EDW based customer purchase data to determine the best projections for customer demand thus combining the power of all three systems (Big Data, Analytical and Traditional) to achieve an optimized supply chain.
Additional data for this analysis is RFID and GPS information from supply chain management (SCM) systems that provide location and quantity of product inventory within the supply chain. Analyzing, storing and filtering this volume of information are again best suited for a Hadoop environment.

Once the data is integrated from these systems the analytical platform can correlate near real-time customer and inventory information to optimize the transportation from warehouses to physical stores via truck or direct to customer via shipping partner.

This use case example utilizes filtered SCM geo-location data and customer sentiment projections from a NoSQL stack along with the inventory and customer history from traditional data management platforms in combination to leverage the complex workload capabilities of an analytic platforms to minimize shipping costs, and maximize profit margins in a way that neither NoSQL or traditional RDBMs are capable of executing alone. By combining the unique power of each of these solutions retailers can realize significant supply chain optimizations resulting in millions of dollars in annual savings and increased revenue.

**Financial Services CRM**

To financial services organizations, a critical component to competitive advantage is development of an effective Customer Relationship Management (CRM) strategy for sales uplift. Businesses that can effectively use customer-centric approaches to cross-sell and up-sell their financial instruments have an enormous leg up on their competition. In retail banking and wealth management, improving cost structure and tailoring products to customer needs rather than driving customers to products is key to the future of financial services organizations. And financial services firms must perform this analysis on a scale and at a speed that is superior to others in their industry while integrating a mix of tens of thousands of products, millions of customers and thousands of points of contact.

To gain the necessary insight these companies are analyzing social media and click stream information stored in Hadoop implementations with externally sourced credit ratings along with current and historical sales information stored in their operational systems in traditional RDBMS platforms to create a responsive customer-centric product recommendation engine. Armed with this analysis, a financial services organization can better match its products with customer needs and customize products to the needs of customers rather than searching out new customers who might fit a particular product. Single digit reduction of sales and marketing expenses at leading financial services companies can equate to millions of dollars in savings per year.

Analyzing all of these desperate data sources on a single platforms isn’t effective. The volume and variety of the data don’t match traditional platform functionality and the level of analytical precision associated with credit scoring and customer lifetime value analysis is inappropriate for NoSQL.

The most effective use of resources is to integrate the correlated and processed customer interaction information from the Hadoop implementation with filtered customer transaction information from the operational systems and external data sources in an analytical platform to perform the recommendation analysis.


The analytical platform provides ability to load the depth of correlated data from the NoSQL environment effectively and quickly. It can also perform analysis at the speed and precision required for this type of customer-centric sales uplift program, in particular for a financial services organization. Finally, it does so at an effective cost in both terms of time to implementation and hard dollar costs since the financial services IT department is not attempting to bend the data to one data management platform or another.

Additional value is added to this scenario as it’s critical for financial services organizations to factor “risk aware” elements into this process. Identifying customer wants and needs is only part of the solution, analyzing risk and implementing this insight into the process is critical. Analyzing multiple data sources spanning social media, click stream analysis, credit ratings sales information and risk models can only been executed on highly proficient analytic platforms designed for the complex workload. Bringing all of these elements together to serve clients, drive revenue and reduce risk is clearly a differentiating value proposition for smart financial services organizations. Delivery of these solutions comes from the optimized use of all three platforms with an analytical platform acting as the bridge between the other two.

**Strategies for Success**

Big Data analytics aren’t confined to NOSQL and New SQL platforms. Additional data is generally needed for final analysis. The ability of analytic platforms to execute on complex workloads positions the solution strategically in the extended data management landscape. To increase performance within your specific data universe consider implementing an analytic platform to add flexibility and speed to your environment.

Expanding data management solutions provide you with flexibility to locate data and workload on the best possible platform for performance. Companies who are attempting to solve these new challenges require a wider set of tools to deliver to analytic stakeholders. It is critical to architect an environment that enables easy data movement between platforms to leverage this strategy and take advantage of having the right platform for processing and storage. The ability to offload data from Big Data platforms to analytic platforms provides the needed flexibility to more efficiently leverage your big data investment. Some sophisticated analytic platforms offer the opportunity to query Big Data without moving it. This adds additional flexibility and power to your solution.

At the core, Big Data isn’t just a challenge addressed by Hadoop and other new Big Data frameworks. Big Data has a home throughout your data ecosystem. The best strategy is to have a flexible array of solutions designed to leverage your overall data investment by ensuring you can use the best possible systems for processing and storage of the data. Analytic platforms enable the flexibility required to fully gain value from Big Data investments.
About Enterprise Management Associates, Inc.

Founded in 1996, Enterprise Management Associates (EMA) is a leading industry analyst firm that provides deep insight across the full spectrum of IT and data management technologies. EMA analysts leverage a unique combination of practical experience, insight into industry best practices, and in-depth knowledge of current and planned vendor solutions to help its clients achieve their goals. Learn more about EMA research, analysis, and consulting services for enterprise line of business users, IT professionals and IT vendors at www.enterprisemanagement.com or blogs.enterprisemanagement.com. You can also follow EMA on Twitter or Facebook.

This report in whole or in part may not be duplicated, reproduced, stored in a retrieval system or retransmitted without prior written permission of Enterprise Management Associates, Inc. All opinions and estimates herein constitute our judgement as of this date and are subject to change without notice. Product names mentioned herein may be trademarks and/or registered trademarks of their respective companies. “EMA” and “Enterprise Management Associates” are trademarks of Enterprise Management Associates, Inc. in the United States and other countries.

©2012 Enterprise Management Associates, Inc. All Rights Reserved. EMA™, ENTERPRISE MANAGEMENT ASSOCIATES®, and the mobius symbol are registered trademarks or common-law trademarks of Enterprise Management Associates, Inc.