Big Data and Business Analytics comes of age

The world of enterprise computing is now starting to see the incorporation of new technologies that have been developed and adopted by the new media companies like Google, Yahoo, Facebook, LinkedIn et al. These technologies include systems like large-scale distributed file systems like Hadoop, which can handle enormously large data sets, cloud computing and virtualization, html 5, smart mobile devices and many more. We are now on the cusp of yet another revolution which will force many industries to change how they process their data and what products and services they can offer.. We are already seeing this in the new media world, where the dramatic rise in the generation and use of social media data and it's obvious value is rapidly changing (once again) how products are marketed and purchased.

The analysis of the data generated by social media companies like Facebook, LinkedIN and the thousands of other related companies have forced companies to adopt new approaches to both gather, store and analyze the data. In addition, successful companies have to be able to scale their systems up and out faster than they can possibly handle on in-house equipment. This in turn has given rise to companies like Rackspace, Amazon and Google to develop architectures that allow the rapid deployment of very large amounts of computing power, essentially allowing companies to outsource all or part of their infrastructure so that they have on demand computing. These changes will have a dramatic impact on both industry and society.

The use of "Business Analytics" will spread rapidly as the knowledge of how to handle and analyze the huge amounts of data now available becomes mainstream. We have been seeing these changes over the last ten years, as first Google and then other companies developed new approaches to handling and analyzing data. These approaches typically involve thousands or tens of thousands of computers that can be harnessed together to attack seemingly insolvable problems. Google, Yahoo and Facebook have datacenters of sizes that would have been beyond comprehension a decade ago. The development of these immense computation centers has forced the computer industry to adapt and develop processors that are much "greener" than the previous generation, i.e. they use much less power and generate much less heat. In addition, these data centers are now being built near power and cooling (i.e rivers and large bodies of water) so that their energy costs are reduced. (Try to imagine Google's electric bill). The efficiency of these new monstrously large data centers means that we are seeing a move back to more centralized computing, where many companies may find it less expensive to outsource some of their computing to a cloud computing provider. This in turn raises a host of security, reliability, contractual and network related issues.

However, Big Data and Business Analytics technologies have now come of age. Within the last few months, both Oracle and EMChave announced NOSQL solutions for unstructured data. This week, Microsoft joined in by announcing that SQL Server 2012 will act as a front end to Hadoop. IBM has been supporting hadoop for years (SEE), and has incorporated it into it's Business Analytics practice. It also used hadoop as the back end for it's Watsonsystem that won at Jeopardy a few months ago.

But the recent announcements by EMC, Oracle and Microsoft will bring these technologies directly to the corporate environment and to many CIOs who have never heard of them.

We are entering a new era. It will be interesting to see how fast these technologies are adopted by major corporations. Some companies, like Oracle and IBM, would seem to have an advantage. For instance, Oracle controls Java (hadoop is written in Java), Lustre (see previous post) and has the ability to sell complete solutions including the hardware and software and consulting, IBM can also offer complete solutions.

The next announcement I expect to hear, is that the major consulting companies will announce practices centered around "big data". This could generate lot's of business for them.

Stay tuned...

Posted 28th October 2011 by Norman White

Labels: big data business analytics cloud computing haoop IBMmicrosoft Oracle Grid Engine Virtualization Watson