Emerging Technologies in Business Intelligence and Advanced Analytics

[1]Nayem Rahman, [2]Dale Rutz, [3]Shameem Akhter, and [4]Fahad Aldhaban

[1]IT Business Intelligence, Intel Corporation, Hillsboro, OR, USA
[2]Supply Chain IT, Intel Corporation, Santa Clara, CA, USA
[3]Department of Computer Science, Western Oregon University, Monmouth, OR, USA
[4]Department of Engineering and Technology Management, Portland State University, Portland, OR, USA

Email: [1]nayem.rahman@intel.com, [2]dale.m.rutz@intel.com, [3]akhtershameem93@yahoo.com, [4]aldhaban@yahoo.com

Abstract—Modern business entities are awash in raw data due to the proliferation of technology. In order to be competitive this data must be refined and organized into trusted information in order to make business decisions in an ever changing environment. Organizations must identify and respond to trends and opportunities on a near real time basis. To survive business organizations are making efforts to improve and enhance their decision making capability using business intelligence (BI). In order to be effective they must improve the performance of BI tools and while driving down associated costs. The rate of data growth is alarming, even a few years ago the hardware and software available made most companies unable to capitalize on the opportunity to harness the power of information within the raw data. Recently, because of the advancement of computing technologies, software engineering, data warehousing technologies, cloud computing, computer processing powers and the emergence of smart-phones the demand for business intelligence tools has increased tremendously. These tools have become more sophisticated than ever before, allowing the perusal and refinement of huge datasets. Emerging technologies and methods are being adopted steadily by business organizations in order to make sense of the explosion of data. This article gives an account of the emerging technologies in business intelligence and data warehousing that significantly improve the performance of data warehouses and consequently, business intelligence tools.

Keywords— Advanced Analytics, Big Data, Business Intelligence, Cloud Computing, Data Warehouse, In-Memory Analytics, Mobile BI

Earlier version, "Emerging Technologies in Business Intelligence" by Rahman, N., Aldhaban, F. and Akhter, S. appeared in Proceedings of PICMET '13. ©2013 IEEE.

Manuscript received on 15 August 2014 and accepted for publication on 29 September 2014.

1 Introduction

TODAY'S business environment is more competitive than ever before due to a number of factors including the global nature of business, the proliferation of technology and the explosion of data in the environment. Supply chains span the globe and business associates must share information and make decisions on a near real time basis. This is made possible through computing technologies [1], software engineering, data warehousing technologies, cloud computing, and lately with the emergence of hand held devices such as tablets and smartphones. The presentation of information on these devices is often used by organizations to make informed business decisions in changing environments.

Companies typically create and collect data in operational data stores. This data is then moved to the enterprise data warehouse for various analytical needs. There is a push to access information faster, often as soon as it is created. To achieve that database engines need to have a faster processing capability [30]. While organizations want a state of the art data warehousing and BI environment they also continuously strive to bring down their IT infrastructure budget. Data warehouse and BI tools involve huge capital expenditure. Small and medium-sized companies cannot afford that. With the advent of cloud computing data warehouse infrastructure costs become incremental to the size of the business need as cloud services allow companies to pay-for-use avoiding the need to own a data warehouse, which can be both costly and time consuming [3] to build. Organizations have started to move from traditional server-based data warehousing to a private cloud.

Businesses are facing challenges in today's environment because of the exponential growth in data year over year. Data is created by business transactions, mobile devices and individuals as well as business entities who share information electronically with each other. Addi-

tionally more and more information is published electronically rather than using more traditional paper methods.

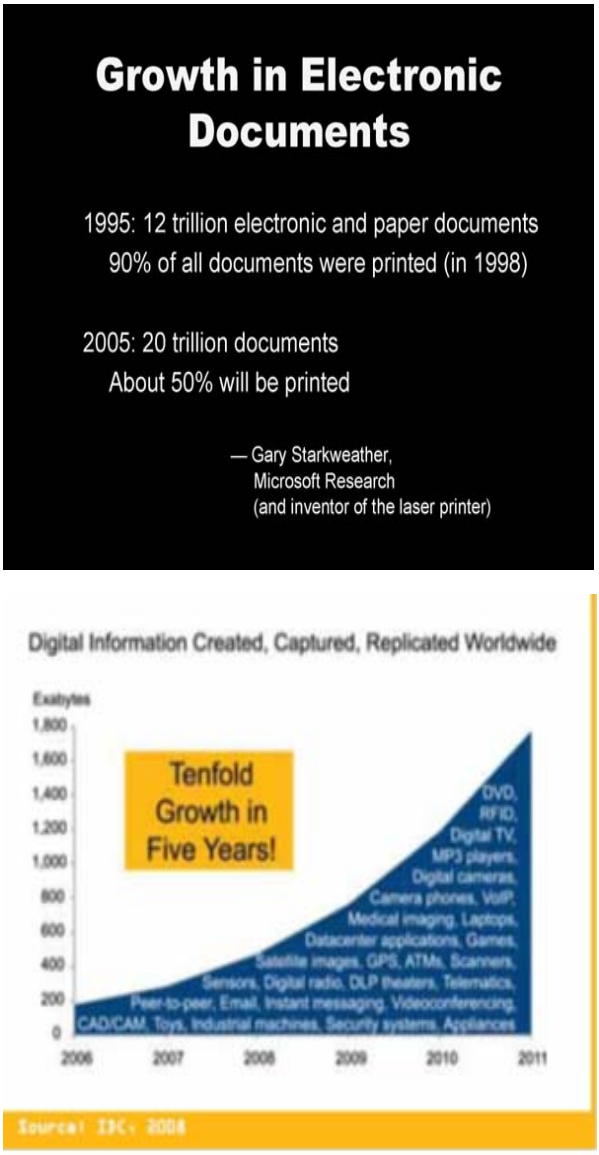

Fig. 1: Growth in Electronic Documents.

Consequently, data storage and retrieval response time increases. So, data warehousing tools and database needs to be more powerful with parallel processing capability to expedite refreshing data warehouses. Fortunately, new tools and parallel processing database engines [40] are now available to assist with overcoming that challenge. Business organizations find business value in unstructured data which are in a huge volume – mostly a few terabytes to hundreds of terabytes. This data comes from many sources including mobile, social media, videos, sensors and surveillance [46] and has a new name called 'big data' as opposed to normal data. Conventional databases are used to store, process, and manage structured data. For big data, which is mostly unstructured and in huge volume, a completely new set of computing technologies have emerged [8]. Business organizations are showing special interest in big data to generate new business out of it. To view, analyze, and visualize both structured and unstructured data BI and visualization tools have emerged. The smart phone and tablets have brought Mobile BI into picture [29], [15].

Business organizations need to deploy emerging technologies in data warehousing (one with a parallel processing capability) and state of the art BI tools (e.g., in-memory analytics capability) to achieve faster data processing of, business intelligence to help making strategic decisions [11], [47] at the right time. To facilitate businesses to make right time decisions a series of new technologies have emerged. In this article, we will cover latest database and business intelligence technologies.

2 Literature Review

In today's data-driven business decision making environment data warehousing and business intelligence play a significant role and are dependent on each other [12]. Data warehousing is deemed one of the six physical capability clusters of IT-infrastructure [51]. Data warehousing has been a research topic for the last two decades [5], [22], [28], [43], [52]. Business intelligence is also gaining growing significance [23], [49] over the last one decade. Data warehouse maintenance [37], implementation [33] and best practices have been explored [53], [30]. Researchers and practitioners have written papers on business intelligence design [20], [39], [9], [38], [54]. BI tools [14], [41], [48] have flourished significantly.

In business organizations the data volume has been increasing significantly every year. Data warehouse users express concern about slowness of access to time-critical data [45]. This has been putting continuous pressure on IT departments to improve performance [25] and efficiency of IT infrastructures. There is a strong correlation between information technology capability and organizational agility [24]. To speed up ETL processing in data warehousing Tank et al. [45] suggest techniques to join operations and data aggregation. Allen & Parsons [2] propose adjusting and reusing existing queries to help improve performance of data warehouse data retrieval process.

Chen et al. [10] state that business intelligence and analytics has emerged as an important area of study to solve data viewing related issues with both 'normal' and big data in business organizations. Watson et al. [50] emphasizes that "to be successful with real-time BI, organizations must overcome both organizational and technical challenges." Business organizations need to adopt emerging technologies to overcome technical and performance issues with traditional tools. The real-time BI helps in making right time business decisions which in turn al-

lows for potential increase of revenues [50]. Ramakrishnan et al. [34] examines how external pressures influence the relationship between an organization's business intelligence (BI) data collection strategy and the purpose for which BI is implemented. Steiger [42] asserts that "BI techniques can be applied to knowledge creation as an enabling technology."

In this article, we show the latest technologies that emerged in data warehousing, BI, big data analytics and in cloud computing, and how they help support faster decision making. These technologies help in decreasing latency in data warehouse refreshes and without impacting performance and resources consumption [31]. These technologies allow for achieving maximum benefits in terms of efficiency, revenue generation and cost avoidance.

3 Parallel Processing DBMS as Emerging Technology

A very few companies have a parallel processing data warehousing architecture. Lately, other DBMS companies have been trying to implement similar kinds of technology. As we see tremendous data growth in medium to large companies, current commercial databases encounter huge amounts of data that need to be processed in loading and business intelligence purposes. Most of the commercial databases are not capable of processing millions of rows within a few seconds to support business intelligence decision making. Parallel processing architecture of DBMS is the right technology towards that endeavor.

In this article, we provide an overview of a parallel processing architecture of a commercial DBMS. Figure 2 shows the parser engine (PE) at the top of hierarchy. The parser of the PE parses and optimizes SQL requests and then dispatches the optimized plan to AMPs over BYNET. After SQL request is processed the query results are returned back to the requesting user via the BYNET. The BYNET loosely couples Symmetric Multiprocessing (SMP) nodes in a multi-node system. There are two network links for each node. The AMPs and PEs send and receive messages using BYNET. It provides communication path among nodes. It merges SQL answer sets back to PE. The BYNET enables parallelism.

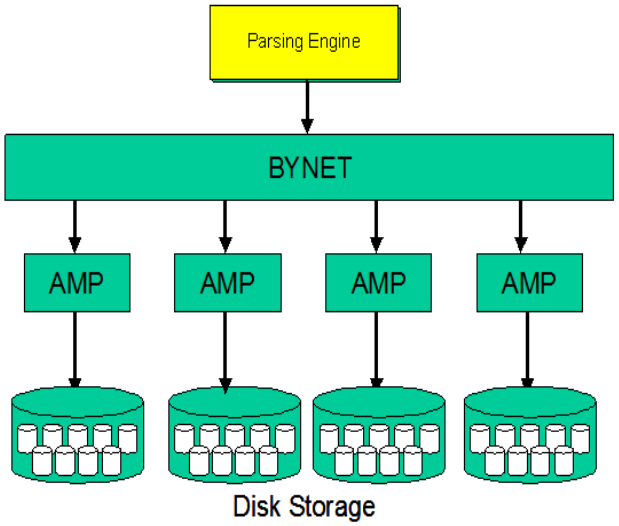

Fig. 2: A Parallel Processing DBMS.

In a parallel processing DBMS architecture a large number of individual access module processors (AMP) are used. The AMPs in DBMS work in parallel. This is also called shared nothing architecture. The data of a table does not get stored into a single AMP and instead is distributed across all AMPs. A hash algorithm decides which record to go to which AMP based on primary index (PI) of the table and during data retrieval all required AMPs participate in the data retrieval process based on a SQL query. That is how a parallel processing DBMS maintains parallelism during both data store and retrieval processes. The data access module processor (AMP) stores and retrieves the distributed data in parallel. In data retrieval process each AMP is only responsible for the rows it stores. A particular AMP cannot pull rows that belong to a different AMP. As far as the AMPs are concerned, it owns all of the rows. The AMPs cannot access each other's data due to shared nothing architecture. The AMPs are designed to work in parallel and return the request results in the shortest possible time.

In SQL writing the ETL programmers must make sure that their SQL takes the parallel processing architecture of DBMS into consideration. Parallel processing database system is a key technology to adopt in handling large volumes of data when loading data warehouses as well as retrieving data from data warehouses for business intelligence purposes using reporting and data mining tools.

4 Emerging Technology in Big Data

Big data refers to dataset, which conventional DBMS cannot store, manage, or analyze. So big data is not limited to analytics [35], instead, big data enables collection, storing, organization and retrieval of information, which in turn enables analyzing and sharing. All of these activities pose great challenges when dealing with large data volumes in a variety of formats. Data comes from many

sources including sensors embedded in processes, systems consisting of operational data, images, videos, documents, mobile devices, and the Internet. Social media is also contributing to the explosion of data in the environment as people post and share information about themselves and their preferences, businesses are quick to capitalize on that data. While organizations have been encountering large volumes of data for years, the exponential growth of data combined with new technologies (e.g. mobile devices and social media) has resulted in a need for better ways to organize and analyze the information. With the maturity of computing technology, processing power, and other data processing and BI tools, organizations are able to take advantage of big data to provide greater insights while assessing new business opportunities resulting in better decision making. Studies show that organizations that adopt data-driven decision making have been successful in increasing productivity 5-6% higher than competition [27]. Big data is used to empower organizations to improve their predictive capabilities [35].

Characteristic	Description	Influencer
Volume	A few terabytes to hundreds of terabytes of data need to be captured, processed, stored, and analyzed	Data volume keeps growing in source
Velocity	Given the volumn the data need to captured, processed, and displayed faster for right time business intelligence and decsion making	Increase in data sources. Improved computing, processing, BI & Visualization technologies
Variety	Includes a variety of data sources with unstructured, semistructured, and structured data. More than 90% unstructured	Sensors, social media sites, digital pictures, video, transaction records, and communication surveillance
Veracity	The quality and provenance of received data. As data most cases data is no structured data consistency is an issue	Data-based decisions require traceability and justification
Value	Provides greater insights generating new business value	Corporate business value

Fig. 3: Big Data with emphasis on 5 V's [46].

Big Data is identified by five factors - volume, velocity, variety, veracity and value. As the name says big data references an enormous amount of data which cannot be handled by conventional database systems and associated tools. Given the volume of data, it has become important to have the capability of receiving, processing and storing data faster (velocity). Big data is also refers to data that comes from many sources and in different formats (variety). The data is unstructured (more than 90%), semi structured and structured. This unstructured nature of data invalidates conventional database systems, which are meant for managing structured data. As data is mostly unstructured data consistency issue comes into picture (veracity). Again, that is due to the nature of source data (mostly unstructured). Business finds value as long as there is traceability and data processing is done by following some processes. With regards to the characteristic, business organizations find business opportunity in big data. One good example is predictive analytics using big data. Big data system provides capabilities for analyzing a greater breadth and depth of data [46].

The big data challenge compels computer scientists, programmers, and information technology professionals to come up with a new paradigm. It is about a complete set of new technology, tools and techniques to receive large volume of data, process them, organize, store and display. There are several technological advancements that have recently occurred which effectively deal with 5 V's of big data. HBase database has emerged as a column oriented database scaling to billions of rows. To handle large volume of data open source framework, Hadoop provides a distributed file system (HDFS). To process large volume of data MapReduce is used for parallel computation on server clusters. Thus big data is distributed and a small portion of data resides on each node. This is shared nothing architecture. Each node possesses its own autonomic unit of CPU, RAM and storage. Not much data movement is needed as the processing occurs where data resides. MapReduce allows lower cost processing of massive data. Hive provides capability of data warehouse with SQL-like access. On the data mining front Mahout provides a library of machine learning and data mining algorithms. Sqoop is used to import data from relational databases. Zookeeper tool is used as a configuration management and coordination. From big data capable database systems there are a couple of new database systems including Hbase and in-memory database. Parallel processing DBMS is best suited for big data because it has parallel processing architecture.

Business organizations have been conducting a variety of test cases using big data to prove usefulness and fulfill business needs. In one of our test cases we faced computing and scalability issue with existing technologies (database and other applications) in comparing all pairs of a few million proteins. This search and comparison activities overburdened the conventional database tables. Later we needed to do this comparison between 20 million proteins. We realized that this cannot be accomplished with existing technology. Here, Big Data technologies come into picture – Hadoop, Map Reduce, Hbase, Hive, etc. Our experiments show that big data technologies has to do the said comparison and analysis by reducing the processing time from days to hours. Another use case was about dealing with medical monitoring data to improve patient outcomes. "The patients routinely are connected to equipment that continuously monitors vital signs, such as blood pressure, heart rate and temperature. The equipment issues an alert when any vital sign goes out of the normal range, prompting hospital staff to take action immediately. The use case result is an early warning that gives caregivers the ability to proactively deal with potential complications, such as detecting infections in premature infants up to 24 hours before they exhibit symptoms" [46]. We conducted another use case to come up with a smart traffic intelligence system, a predictive analytics using Hadoop technologies. The challenge was to analyze city traffic data to derive statistics for crime preven-

tion, information sharing, and predictive traffic analysis. By using real time traffic data predictive analytics was performed. The big data technologies helped with generating automated queries for traffic violation, data mining of fake licenses in a minute based on data captured for a week. This has improved the predictive traffic forecasting capability of city authority.

Chandramouly & Stinson [8] provide an architectural overview of big data solution that allows for effective use of BI tools to run business with operational efficiency and competitive advantage. While big data holds hidden information, business analysts need to read or view that business information using some analytical and visualization tools [21]. Heer and Kandel [17] suggest some interactive analysis tools for big data to empower data analysts to formulate and assess hypothesis in a rapid manner. To make meaningful information from big data, visualization matters. Visualization of data helps us to understand data, see patterns, spot trends and detect outliers [16]. Heer and Shneiderman [18] emphasize several visualization techniques of big data. They propose a taxonomy of interactive dynamics for visual analysis consisting of data view specification, view manipulation and processes.

5 In-Memory Analytics in Business Intelligence and DBMS

In-memory data processing is a very new technology that has recently emerged. It's been used in both database and business intelligence space. In-memory has several key performance benefits [13]: dramatic performance improvements; cost-effective alternative to data warehouses; discover new insights; and connect insight with action.

Read and Write Capabilities
Centrally Managed Data, Business Hierarchies, Rules and Calculations
Empower Business Users to Analyze any Combination of Data
High Impact Visualizations
Extend and Transform Excel
Designed for Modern 64 bit Architectures
Easy to Insall and Easy to Use

Fig. 4: Capabilities of In-Memory Analytics [13].

One leading BI company [13] provides several capabilities with in-memory analytics. These include read and write capabilities; centrally managed data, business hierarchies, rules and calculations; empower business users to analyze any combination of data; high impact visualizations; extend and transform excel; designed for modern 64 bit architectures; and easy to install and easy to use. A leading ERP company has come up with in-memory database technology which provides better performance of analytics and transactional applications. The in-memory database is being positioned to handle big data in terms of several terabytes of data in memory for analytical purposes.

Some leading commercial database companies have launched its in-memory database and business analytics technology. Their in-Memory machine features an optimized BI foundation suite and its 'TimesTen' In-Memory database. The BI Foundation takes "advantage of large memory, processors, concurrency, storage, networking, kernel, and system configuration of the exalytics hardware. This optimization results in better query responsiveness, higher user scalability" [27].

6 Data Warehousing and BI with Cloud Computing

Data warehousing projects are very expensive. Normally medium to large companies maintain their own data warehouse. On the other hand, companies of all sizes have data growing over the years. Data warehouse and BI on the cloud have opened the door for all sizes of companies to use these technologies. The Cloud is transforming the economics of BI and opening up many new possibilities for organizations of all sizes [47]. With cloud, business organizations will find it relatively easy to fund data warehousing projects given the low cost and no maintenance involved. For cloud-based data warehousing long term capital expenditures is not needed. Businesses can pay for cloud service on a weekly, monthly or pay per service basis. Cloud data warehousing and BI will allow business organizations to conduct more short-term ad-hoc analysis. Building an organization's own data warehouse takes a long time to set up infrastructure and line up resources. With cloud individual business organizations do not have to worry about infrastructure and logistics. With cloud technology it takes a few hours or days to get an initial data warehouse created. Cloud-based data warehouse is economically suitable for sandbox kind of development and testing as well as short-lived projects. Cloud-based analytic databases will enable small companies to warehouse and analyze a large volume of data even though their BI budgets and staff are much smaller than larger enterprises. On the other hand, analytic SaaS market will develop faster [19]. Amazon is the leading provider of cloud services. All leading commercial database companies have teamed up with a leading cloud provider to run their databases on the cloud platform (Amazon EC2).

Cloud service providers provide both public and private cloud services. Given that most of the business organizations have financial and other mission critical data, using public cloud is not considered the best choice as public cloud is shared by multiple parties sharing the same network, server, software and hardware. Data security is the biggest concern. While cloud provides financial advantage, concern about security data and restricting access to limited individuals grows as data is stored outside company firewalls on external clouds, therefore industry experts often suggest using private cloud from

the standpoint of data security. They also suggest using multiple layers of data protection (e.g., password encryption) and the higher levels of security [4] to ensure compliance requirements of individual business organizations. Another limiting factor of external public clouds is the difficulty in transferring large volumes of data over the internet to external clouds for storage. Many small companies today which are forced to use public clouds for financial reasons are resorting to shipping hard drives full of data via overnight shipping companies in order to transfer huge volumes of data to the cloud.

As emerging technology, a leading cloud provider recently announced a new cloud-based data warehousing and BI tool [7]. This service provides service in big data analytics as well. The customers can pay per service basis to make it affordable to small and medium-sized businesses [26]. The data warehouse companies are already on the market [4] with cloud-based data warehouse appliance offering. So with the offering of cloud-based data warehousing and BI technology by these leading companies, the adoption of data warehousing and BI should expand much faster [19].

It is clear that with the advent of big data, business professionals need better, more efficient ways of consuming that data. The goal is to use the information gleaned from big data to make better business decisions faster, thus improving velocity and ultimately, profitability.

7 Advanced Analytics

The buzz over advanced Analytics in the business world came about when companies started showing off dashboards full of visually appealing controls and splashy colors. While the initial reaction was one of excitement concerns over form and function soon surfaced. Expertise in creating and maintaining these dashboards was scarce and expensive. Further, many of the early advanced analytics dashboards were little more than flashy displays which did little to enhance business decisions. As a result usage metrics were upside down, with initial metrics showing many uses per day, trending down to a point where 6 months later no one was viewing the dashboard at all. Clearly, many early advanced analytics implementations where the result of marketing hype and failed to deliver the promised business value.

Today, many people are starting to use the terms advanced analytics and big data interchangeably. This is because the realization has set in that the goal of advanced analytics is, and should be, to make better business decisions faster. In order to achieve this goal two things are required, those being: (1) Timely access to relevant data which has been cleansed, processed and formatted for rapid access. (2) A presentation layer which allows business professionals to, at a glance, identifies patterns and trends in order to react with appropriate decisions.

In a simple example at a recent conference a presentation layer company displayed a 20 x 20 matrix of characters and asked participants to count the number of letter ‘a’ s. On the next slide participants were asked to do the same thing, with the letter ‘a’s highlighted in red. Consider the tables below and the difference is obvious:

me8xwask;ldjaflksdjf	me8xwask;ldjaflksdjf
yewsghpoe9chefgk;d	yewsghpoe9chefgk;d
ps;kdjgls;adjkglsdjfla	ps;kdjgls;adjkglsdjfla
x;skdjf;lsdkf;lasdkfsl	x;skdjf;lsdkf;lasdkfsl
asldkfjlskdjfdslfkjsadc	asldkfjlskdjfdslfkjsadc

Fig. 5: Matrix of Characters.

So we can think of advanced analytics as being enabled by big data technologies and similarly we realize that big data is dependent on advanced analytics in order for business professionals to be able to consume and apply the information that big data provide. The goal of analytics is to traverse huge amount of data, seeking patterns which can be used to develop business opportunities. Previously this kind of work was done by statisticians using data mining tools. Today there is a proliferation of tools which allow trained professionals to sample, chart, analyze and report on data. This has given rise to a new type of job, the data scientist. These professionals are charged with the responsibility of creating structure and order out of the chaos that can accompany huge volumes of data.

Data Scientists organize data using rules. They define data quality and implement rules which cleanse data. They also set prescriptive guidance such that data falling outside of defined boundaries trigger alters which cause people to investigate. Initial application of these rules has fallen on fraud detection and remediation. It is clear that there are numerous uses of these actives however. Shopping basket analysis is another clear winner here. Once it is noticed that people or businesses who buy product X often are interested in product Y then directed marketing can target individuals who are most likely to consume a product or service.

Shopping basket analysis examples illustrate a shift in the focus of business intelligence. In the past most forms of analytics were backward looking. What happened in the recent past and how does that compare happenings farther in the past and so on. With the advances in hardware and big data software business users can now look at patterns and trends as they occur, giving rise to real time business intelligence. This allows businesses to use techniques such as Predictive Churn analysis to identify vulnerable customers before they move to a competitor’s products or services and win them back before they are even gone.

One of the most important outcomes of the convergence of big data and advanced analytics is that big data can reduce or even eliminate the need to use “sampling”

when doing data mining. Sampling involves using only a portion of the data to analyze and predict and results in a type of error called bias. With big data technology coupled with hardware and software performance improvements data mining can now be run against entire data sets rather than a sample resulting in much more accurate prescriptive analytics and therefore the best possible business decisions.

Many companies in the market have taken advantage of new advances in technology to allow two phases of data warehousing using Big Data. One is called the infinite store, typically a noSQLdatabase which uses file system storage. This gives virtually unlimited capacity at low cost. Companies are coupling this infinite store with and "instant store" which holds relevant information in memory. By using solid state devices, enhanced search and retrieval algorithms as well as the latest processor technologies retrievals which previously took 20 to 30 minutes can be accomplished in less than a second. These advances in technology are allowing business people to make data driven decisions on a near real time basis.

The typical business user accesses information in order to make informed decisions. It is critical that the volume of data which reaches the end users device is reduced to what is needed and relevant. One of the tenants of big data is to host not only detail but also aggregate, or processed data. Detailed data is processed into aggregates and those aggregates are what get transmitted to the presentation layer. Thus a salesman can pull out a phone or a tablet and select a customer and get information such as year to date sales by product category and time between orders as they reach the customer site. They can then use that information to suggest to the customer that they may be nearly out of certain items which have not been ordered recently. Similarly they might compare this customer's year to date purchases with that of last years and notice a change in pattern which may trigger them to make suggestions or ask questions. Placing these kinds of data tools into the hands of the sales force is necessary in today's business climate.

8 Presentation Layer Form Factors

Modern technology allows information to be presented and consumed on a myriad of devices. This is a stark contrast to a world even 5 years ago where almost all business data was presented and consumed on computers. Clearly today this is not the case. In today's environment people are using phones as "pocket computers." They are using tablets, notebook computers, desktop computers, and displays in their cars and even television monitors to access information as well as entertainment. In fact, with the advent of social media the line between data and entertainment has blurred. Where once we had a clear distinction between computers used for viewing columns and rows of data and televisions used for viewing moving pictures today we find a multitude of devices which can be used for either purpose at any given point in time. Further, many of today's "visualizations" involve animation (e.g. bubble charts). Data sources are omnipresent; everything from intranets to the internet to the cloud puts petabytes of data within reach of us all. Companies have responded to the opportunities provided by the onslaught of new devices by releasing software which extends their Advanced Analytics presentation layer software to mobile devices. Today, many companies offer options which will deploy the same interface to computers, tablets and phones. Thus traditional users and mobile users can access the same information and view and interact with it in the same format regardless of their location or the device at hand.

This evolution of devices has led to a corresponding revolution in software engineering. Requirements today are very different than they were even a few years ago. Hand held devices are often more powerful than computers of yesterday. As a result computers are used less and less, while devices such as tables are omnipresent, at conferences, in the office, in the home and in the hands of the sales force. Users require interfaces such that the behavior of the interface is largely the same regardless of the device being used. This requirement is answered with something called the "app store." Each major device manufacturer has an app store. Similarly companies have internal app stores for their proprietary applications. Developers build applications in provider agnostic languages such as Java and HTML5. They can then upload the raw application to different app stores for different devices and they will be compiled and work correctly across applications. These same applications can also usually be hosted on a web page. Thus users have access to the same data in the same format on their lap top, their phone, their tablet, their desktop computer and also often their gaming device as well as their television.

With the proliferation of devices come increased security concerns. Hand held devices are designed to be able to broadcast information using technologies such as blue tooth because they are communication devices. Therefore applications which are secure on the web may or may not be on a hand held device. Also differences between device manufactures mean different app stores which results in the same application being more secured on some devices as compared to others. Users are concerned and with good reason. Businesses are even more alarmed because not only is the information transmitted to or from a device over the airwaves less secure but the devices themselves are easily, and therefore frequently, lost or stolen. A whole host of products and services are offered on the market to address these information security concerns. These concerns do present a barrier to entry when it comes to sensitive organizational data.

9 Business Intelligence in Mobile

Recent advancements in mobile devices have had a

positive impact in the business intelligence field. Industry surveys show that mobile BI is gaining popularity. According to a recent survey, "24% of enterprises already use or are piloting mobile BI applications, while 37% are considering mobile BI for near-term implementations" [44]. In another survey it was observed that "87% of the respondents reported that they planned to use a mobile device to help make purchasing decisions during holiday shopping" [29]. The trend of computer-based reports is changing with the emergence of Business Intelligence system applications that run on mobile devices, such as smart-phones and tablet computers [6]. The Gartner research findings indicate that one third of BI functionality will be consumed through hand-held devices. With the growing popularity of the tablet, the adoption rate of mobile BI is expected to soar [36]. The proliferation of apps which run on multiple devices (phone, tablets computers) has allowed access to BI information at anywhere and anytime. These devices have downloading capabilities faster than expected due to the proliferation of cloud based technologies. Additionally, touch-screens have improved the user experience [19]. All these speak for business intelligence in the Mobile era as opposed to desktop-based business intelligence. The corresponding increase in mobile data traffic based on the explosion of BI on mobile devices is startling:

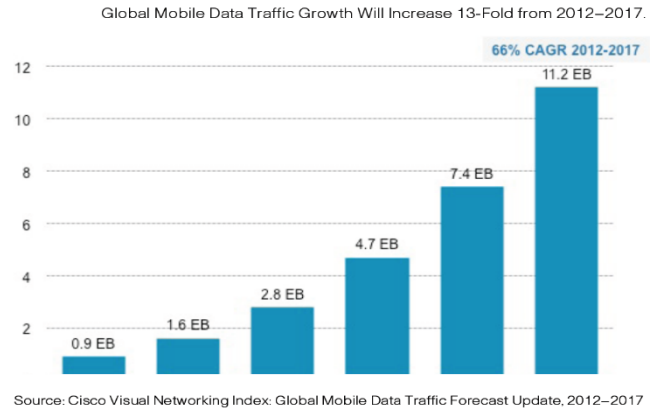

Fig. 6: Global Mobile Data Traffic Forecast.

Robb [36] reports that the top ten mobile intelligence apps. All of this application operates on iPhone, iPad, Blackberry, and Windows Mobile. Each offers several great features. Some of these mobile applications have the ability to navigate through displays with interactive graphs, tables and charts. Some other mobile architecture provides data visualization to mobile devices. Some companies offer bundled server software that integrates with most BI platforms and data warehouses. Some of the BI companies' Mobile architecture extends graphs, grids, enterprise reports and information dashboards and its features also include out-of-the-box integration with Google maps [36].

With the advent of smart phone and tablets the Mobile Decision Support System (MDSS) has emerged as a new decision support system in this early 21st century. Haghighi suggests that MDSS can be beneficial to application domains where critical decisions need to be made under time pressure and the decision-makers are on the move [15]. There are many industries that use the emerging mobile BI technologies. These include mobile healthcare, emergency management, mobile commerce, education, mobile banking [15], purchase and selection decisions, and negotiations [29].

Many of these applications are game changers for the industries they serve. For example health care has traditionally been provided by doctors meeting face to face with patients who travel whatever distance is necessary to see the doctor. Today, companies such as Intel Corporation manufacture devices and suites of applications which allow medical technicians to meet with patients in remote settings. Technology is used to connect the remote setting with the doctor who is supplied with all of the information collected while simultaneous interacting with the patient. The doctor can meet face to face with the patient while at the same timing use real time analytics to identify issues and potential resolutions.

Another example is education. Previously most academic materials were available only in printed media. Today we are seeing a major shift, especially with emerging nations providing content only in electronic form as printed books become obsolete due to expense and environmental footprint. This electronic media lends itself to automated search and retrieval, meaning students of tomorrow will have limitless access to information and will need better and better tools to distill data down to relevant information. In addition, they will be accessing this electronic content more and more from alternative devices (eBook readers, tablets, their cars and even phones).

Traditional sales and service providers are also seeing their industries shift to embrace advanced analytics and near real time business intelligence. Take for example the providers of in home video content. Their predictive churn models can identify customers who fit into various categories of users who are likely to leave, allowing companies to reach out to these customers to offer them bundles of services at discounted prices which tempt the user to stay with the provider. Other algorithms can determine the best way to reach out to the customer (printed materials, e-mail, a phone call, a text alert, et al).

What if-scenario (e.g. write-back) are quickly becoming an integral part of advanced analytics. Data retrieved from data warehouses contains the actual facts. Often business people need to know what the result would be if key facts changed. Write-back analytics allow users to enter new facts in a what-if scenario and then rerun models to determine the impact of the change. There is a need to save not only the parameters submitted for the what-if analysis but also the resultant model. Many companies are offering this type of functionality out of the box.

10 Concluding Remarks

In this article we explored the emerging technologies in data storage, retrieval as well as business intelligence and advanced analytics. We also discussed the benefits of using these new technologies. These emerging technologies and methods are and will continue to be steadily adopted and enhanced in the near future by business organizations. Use of these new technologies in data warehousing and business intelligence will help organizations make strategic and tactical decisions at the right time and increase revenues. Additionally these technologies are required in the existing market in order for companies to remain competitive. Access to relevant and understandable information in our ever changing environment is clearly the only way to survive.

Acknowledgment

Earlier version of this article was presented at the IEEE Portland International Center for Management of Engineering and Technology (PICMET 2013) Conference, San Jose, California, USA, July 28 - August 1, 2013 [32]. This article has been substantially revised and enhanced, based on comments received during conference presentation and comments received from anonymous reviewers of this journal. The author would like to thank the anonymous reviewers for their constructive comments provided on earlier drafts of this article.

References

[1] S. Akhter, N. Rahman, and M.N. Rahman, "Competitive Strategies in the Computer Industry," International Journal of Technology Diffusion (IJTD), Vol. 5, No. 1, pp. 73-88, January-March, 2014.

[2] G. Allen and J. Parsons, "Is Query Reuse Potentially Harmful? Anchoring and Adjustment in Adapting Existing Database Queries," Information Systems Research, Vol. 21, No. 1, pp. 56-77, 2010.

[3] R. Armstrong, "Data Warehousing: Dealing with the Growing Pains," IEEE Proceedings, 1063-6382/97, pp. 199-205, 1997.

[4] J. Bair, "Emerging Cloud Appliances for Business Intelligence and Data Warehousing: 6 Challenges and Opportunities," Retrieved on 04/24/2013 from: http://www.b-eye-network.com/view/16915, 2013.

[5] S. Brobst, M. McIntire and E. Rado, "Agile Data Warehousing with Integrated Sandboxing," Business Intelligence Journal, Vol. 13, No. 1, 2008.

[6] Business Intelligence Strategy, "Smarter Mobile Devices Drive Demand for Mobile Business Intelligence Applications," Retrieved on 01/12/2013 from: http://www.businessintelligencestrategy.com.au/news-views/mobile-business-intelligence/, 2013.

[7] B. Butler, "Is the cloud the right place for your data warehouse?," Network World, Retrieved on 04/24/2013 from: http://www.networkworld.com/news/2012/121012-aws-data-warehouse-264957.html, 2012.

[8] A. Chandramouly and K. Stinson, "Enabling Big Data Solutions with Centralized Data Management," Intel IT White Paper, January 2013, Available at: www.intel.com/it, 2013.

[9] S. Chaudhuri, U. Dayal and V. Narasayya, "An Overview of Business Intelligence Technology," Communications of the ACM, Vol. 54, No. 8, pp. 88-98, 2011.

[10] H. Chen, R.H.L. Chiang and V.C. Storey, "Business Intelligence and Analytics: From Big Data to Big Impact," MIS Quarterly, Vol. 36, No. 4, pp. 1165-1188, 2012.

[11] B.L. Cooper, H.J. Watson, B.H. Wixom and D.L. Goodhue, "Data Warehousing Supports Corporate Strategy at First American Corporation," MIS Quarterly, Vol. 24, No. 4, pp. 547-567, 2000.

[12] C. Eden and V. Padmanabhan, "Building an Enterprise Data Warehouse and Business Intelligence Solution," Intel Information Technology White Paper, pp. 1-11, 2006.

[13] Gartner, Inc., "In-Memory Analytics: Leveraging Emerging Technologies for Business Intelligence," http://download.boulder.ibm.com/ibmdl/pub/software/data/sw-library/cognos/pdfs/ar_inmemory_analytics_leveraging_emerging_technologies_for_business_intelligence.pdf, 2009.

[14] J. Hagerty, R.L. Sallam and J. Richardson, "Magic Quadrant for Business Intelligence Platforms", Gartner, Inc., ID: G00225500, 2012.

[15] P.D. Haghighi, "The New Era of Mobile Decision Support Systems," Journal of Decision Systems, Vol. 22, No. 1, pp. 1-3, 2013.

[16] J. Heer, M. Bostock and V. Ogievetsky, "A Tour Through the Visualization Zoo," Communications of the ACM, Vol. 53, No. 6, pp. 59-67. doi :10.1145/1743546.1743567, 2010.

[17] J. Heer and S. Kandel, "Interactive analysis of big data," ACM Crossroads Vol. 19, No. 1, pp. 50-54, 2012.

[18] J. Heer and B. Shneiderman, "Interactive Dynamics for Visual Analysis," Communications of the ACM, Vol. 55, No. 4, pp. 45-54. doi:10.1145/2133806.2133821, 2012.

[19] Jaspersoft, "Seven Trends That Will Change Business Intelligence as We Know It," Retrieved on 01/12/2013 from: http://resources.idgenterprise.com/original/AST-0043754_Jaspersoft_eBook.pdf, 2011.

[20] Z. Jourdan, R.K. Rainer and T.E. Marshall, "Business Intelligence: An Analysis of the Literature," Information Systems Management, Vol., 25, pp. 121–131, 2008.

[21] S. Kandel, A. Paepcke, J.M. Hellerstein and J. Heer, "Enterprise Data Analysis and Visualization: An Interview Study," IEEE Trans. Vis. Comput. Graph, Vol. 18, No. 12, pp. 2917-2926, 2012.

[22] K. Lam and V.C. Lee, "On Consistent Reading of Entire Databases," IEEE Transactions on Knowledge and Data Engineering, Vol. 18, No. 4, 2006.

[23] A. Lonnqvist and V. Pirttimaki, "The Measurement of Business Intelligence," Information Systems Management, Vol. 23, No. 1, p.32, 2006.

[24] Y. Lu and K. Ramamurthy, "Understanding the Link between Information Technology Capability and Organizational Agility: An Empirical Examination," MIS Quarterly, Vol. 35, No. 4, pp. 931-954, 2011.

[25] S. Mithas, N. Ramasubbu and V. Sambamurthy, "How Information Management Capability Influences Firm Performance,"

MIS Quarterly, Vol. 35, No. 1, pp. 237-256, 2011.

[26] G. Narcisi, "AWS introduces Redshift 'big data' and business intelligence service," Retrieved on 04/24/2013 from: http://searchcloudprovider.techtarget.com/news/2240173984/AWS-introduces-Redshift-big-data-and-business-intelligence-service, 2012.

[27] Oracle Corporation, "Oracle Exalytics In-Memory Machine: A Brief Introduction," An Oracle White Paper, 2011.

[28] F. Payton and R. Handfield, "Strategies for Data Warehousing," MIT Sloan Management Review, 2004.

[29] D.J. Power, "Mobile Decision Support and Business Intelligence: An Overview," Journal of Decision Systems, Vol. 22, No. 1, pp. 4-9, 2013.

[30] N. Rahman, "Refreshing Data Warehouses with Near Real-Time Updates," Journal of Computer Information Systems, pp. 71-80, 2007.

[31] N. Rahman, "Measuring Performance for Data Warehouses - A Balanced Scorecard Approach," International Journal of Computer and Information Technology (IJCIT). Vol. 4, No 1, 2013, Pages, 1-7.

[32] N. Rahman, F. Aldhaban and S. Akhter, "Emerging Technologies in Business Intelligence," In Proceedings of the IEEE Portland International Center for Management of Engineering and Technology (PICMET 2013) Conference, San Jose, California, USA, July 28 - August 1, 2013, pp. 542-547.

[33] N. Rahman, "Temporal Data Update Methodologies for Data Warehousing," Journal of the Southern Association for Information Systems (JSAIS), Vol. 2, No. 1, pp. 25-41, Summer 2014. DOI: http://dx.doi.org/10.3998/jsais.11880084.0002.103.

[34] T. Ramakrishnan, M.C. Jones and A. Sidorova, "Factors influencing business intelligence (BI) data collection strategies: An empirical investigation," Decision Support Systems, Vol. 52, 2012, pp. 486–496.

[35] M. Ricknäs, "Big data not just about the analytics," Retrieved on 04/16/2013: http://www.computerworld.com/s/article/9224986/Big_data_not_just_about_the_analytics_says_Amazon_CTO, 2012.

[36] D. Robb, "Ten Great Mobile Business Intelligence Apps", http://www.enterpriseappstoday.com/business-intelligence/ten-great-mobile-business-intelligence-apps-1.html, 2011.

[37] E.A. Rundensteiner, A. Koeller and X. Zhang, "Maintaining Data Warehouses over Changing Information Sources," Communications of the ACM, Vol. 43, No. 6, pp. 57-62, 2000.

[38] D. Rutz, T. NelaKanti and N. Rahman, "Practical Implications of Real Time Business Intelligence," Journal of Computing and Information Technology (CIT), Vol. 20, No. 4, 2012.

[39] D.I. Sandu, "Operational and real-time Business Intelligence," Revista Informatica Economică, Vol. 3, No. 47, pp. 33-36, 2008.

[40] A. Sen and A.P. Sinha, "A Comparison of Data Warehousing Methodologies," Communications of the ACM, Vol. 48, No. 3, pp. 79-84, 2005.

[41] D. Spy, "5 Top Business Intelligence Tool Vendors According to IDC," Retrieved on 01/12/2013 from: http://www.enterprise-dashboard.com/top-5-business-intelligence-tool-vendors-2007/, 2007.

[42] D.M. Steiger, "Decision Support as Knowledge Creation: A Business Intelligence Design Theory," International Journal of Business Intelligence Research (IJBIR), Vol. 1, No. 1, pp. 29-47, 2010.

[43] V.C. Storey and R.C. Goldstein, "Knowledge-Based Approaches to Database Design," MIS Quarterly, Vol. 17, No. 1, pp. 25-46, 1993.

[44] S. Tabbitt, "Mobile Business Intelligence: Here At Last?" Retrieved on 04/25/2013 from: http://www.informationweek.com/software/business-intelligence/mobile-business-intelligence-here-at-las/240146333, 2013.

[45] D.M. Tank, A. Ganatra, Y.P. Kosta and C.K. Bhensdadia, "Speeding ETL Processing in Data Warehouses Using High-Performance Joins For Changed Data Capture (CDC)," 2010 International Conference on Advances in Recent Technologies in Communication and Computing, DOI: 10.1109/ARTCom.2010.63, pp. 365-368, 2010.

[46] TechAmerican Foundation, "Demystifying Big Data: A Practical Guide To Transforming the Business of Government," Retrieved on 04/16/2013 from: http://www.techamerica.org/Docs/fileManager.cfm?f=techamerica-bigdatareport-final.pdf, 2012.

[47] Vertica Systems Inc., "Transforming the Economics of Data Warehousing with Cloud Computing," Retrieved on 01/20/2013 from http://www.vertica.com/wp-content/uploads/2011/01/CloudTransformsEconomicsOfDataWarehousing-Vertica.pdf, 2008.

[48] D. Vesset, "Competitive Analysis: Worldwide Business Intelligence Tools 2010," IDC Analyze the Future, Retrieved on February 4, 2012 from: http://www.sas.com/news/analysts/103115_0611.pdf, 2011.

[49] H.J. Watson, "Tutorial: Business Intelligence – Past, Present, and Future," Communications of the Association for Information Systems, Vol. 25, No., p. 39, 2009.

[50] H.J. Watson, B.H. Wixom, J.A. Hoffer, R. Anderson-Lehman and A.M. Reynolds, "Real-Time Business Intelligence: Best Practices at Continental Airlines," Information Systems Management, Vol. 23, No. 1, pp.7-18, DOI: 10.1201/1078.10580530/45769.23.1.20061201/91768.2, 2006.

[51] W. Weill, M. Subramani and M. Broadbent, "Building IT Infrastructure for Strategic Agility," MIT Sloan Management Review, 2012.

[52] J. Widom, "Research Problems in Data Warehousing," In Proceedings of the 4th Int'l Conference on Information and Knowledge Management (CIKM), 1995.

[53] B.H. Wixom and H.J. Watson, "An Empirical Investigation of the Factors Affecting Data Warehousing Success," MIS Quarterly, Vol. 25, No. 1, pp. 17-44, 2001.

[54] I. Yermish, V. Miori, J. Yi, R. Malhotra and R. Klimberg, "Business Plus Intelligence Plus Technology Equals Business Intelligence," International Journal of Business Intelligence Research (IJBIR), Vol. 1, No. 1, pp. 48-63, 2010.

Nayem Rahman is a Senior Enterprise Application Developer in IT Business Intelligence (BI), Intel Corporation. He has implemented several large projects using data warehousing technology for Intel's

mission critical enterprise DSS platforms and solutions. He is currently pursuing a Ph.D. in the Department of Engineering and Technology Management, Portland State University, USA. He holds an M.S. in Systems Science (Computer Modeling & Simulation) from Portland State University, Oregon, USA and an MBA in Management Information Systems (MIS), Project Management, and Marketing from Wright State University, Ohio, USA. His most recent publications appeared in Proceedings of the IEEE 26th Canadian Conference of Electrical and Computer Engineering (CCECE 2013) and the International Journal of Computer and Information Technology (IJCIT). His principal research interests include Big Data Analytics, Active Data Warehousing, Data Mining for Business Intelligence, Intelligent Data Understanding using Simulation, and Simulation-based Decision Support System (DSS).

Dale Rutz is a BI Solution Architect in Supply Chain IT. She has been developing software at Intel Corporation for 23 years. Her work has involved all facets of software development including ETL, integration and presentation layer using technologies such as SQL, JAVA, UNIX, C, XML and various proprietary protocols. Ms. Rutz holds an MBA from Santa Clara University and a BS from San Jose State University in MIS and Math. She completed the APICS CSCP (certified supply chain professional) in December 2008. Her focus is on harnessing the power of information.

Shameem Akhter is an Information Technology (IT) Professional. She holds an M.S. in Management and Information Systems (MIS) from Western Oregon University, USA. She had been selected by Western Oregon University (WOU) faculty and staff committee for the Who's Who among Students in American Universities and Colleges recognition for 2011-2012. The committee selected her from a nominated pool of over 900 students to be recognized for her outstanding leadership, service, and scholarship at WOU. Her most recent publications on IT sustainability and data warehousing appeared in the International Journal of Technology Management & Sustainable Development (IJTMSD) and the Journal of Computing and Information Technology (CIT) respectively. Her research interest includes Database Systems, Data Warehousing, Decision Support System, and Information Systems Implementation.

Fahad Aldhaban is pursuing his Ph.D. in the Department of Engineering and Technology Management, Portland State University, USA. He holds an M.S. in Engineering and Technology Management from Portland State University, Oregon, USA, an M.S. in Management and Information System (MIS) from Fairleigh Dickinson University, New Jersey, USA and a B.S. from Al Iman Muhammad Ibn Saud University, Saudi Arabia. He worked as a section manger in Saudi Telecommunications Company (STC) for five years. He published a number of papers including, Exploring the adoption of smartphone technology: literature review', 'Adoption & evaluation of personal health record (PHR) system' and 'Exploring the adoption of smartphone technology: literature review'. His most recent publication appeared in Proceedings of the IEEE Portland International Center for Management of Engineering and Technology (PICMET 2013) Conference. His research interest includes Technology Assessment, Technology Adoption, Quantitative Data Analysis, and Technology Roadmaping.