APPLICATIONS OF BIG DATA ANALYTICS IN ENTERPRISE

Gayatri Hegde[1], Madhuri Rao[2]
Department of Computer Engineering,
Thadomal Sahani Engineering College, Bandra, Mumbai, India
Email:ghegde@mes.ac.in[1], my_rao@yahoo.com[2]

Abstract
The key to success for any enterprise lies in the ability to collect right data at right time and deliver the right data, at the right velocity, and in the right quantities to a wide set of analytics Here in this paper we look into how many big enterprises used big data to improve their efficiency and hence the profit.
Index Terms: Big Data Applications, Big Enterpise.

I. INTRODUCTION

Big data analytics is used to analyze large amount of data sets that contain data of different variety, velocity, and veracity collected from many different channels to find hidden patterns, relations, understand customer preferences and other useful information. There are many big enterprises who have led the example of use of big data analytics. Here we present few findings of few big enterprises who are leaders in use of big data analytics.

A. Big Data at UPS

UPS United Package Service is one of the biggest package shipping company. They gather about 16 petabytes of data[6]. They keep track of drivers and package movement and also keep track of transactions.

They used the system called ORION (On-Road Integrated Optimization and Navigation) algorithm, for data analysis.

UPS capture and track their package movements and transactions. They track about 16 million packs per day for about 8 million customers. Thus they store about 16 petabytes of data per day.

It uses telematics sensors in their vehicles. The data thus collected may include their speed, direction, drive, performance. This data is used to analyse performance and also analyse optimal route structure.

They use ORION algorithm. It is used to analyse 20000 possible routes for each track in real time.

By using big data, they are able to improve the performance of each route and delivery of packages. They saved upto 8.4 million gallons of fuel by cutting 85 millions of miles. With this they could save upto 25 million dollar per year.

B. Big Dataat Carsars Entertainment

Ceasars (formarly called Harrah's) is one of the world's largest gaming and entertainment companies. They run the company with multiple brands and operate in four different continents[7].

Caesars mainly focus on its customers loyalty, marketing and services provided by the company. The objective of the company in using big data is to respond to the customer requirement and provide them required service in real time.

They gathered their customer data from its Total Reward Loyalty Program, web clickstreams, and while playing at slot machine or in the resort.

The company found it traditionally difficult to analyse and react to the customers in real time. Hence they switched to big data. With video analytics in big data tools, they were able to recognize more service issues. Caesar is also now analysing mobile data so that it can keep their loyal customers and also attract less frequent customers with their real time offers.

Caesars use both Hadoop clusters and open source and commercial analytic software for their analysis.

C. Big Data at United Health Care

United Healthcare provides health services to more than 50 million people. They have around 8.5 million physicians and more than 6000 hospitals associated with them. The big problem faced by them was in ensuring correct claims and if found correct than paid on time. They deal with around 10 TB of data daily[7].

United Healthcare is now able to use Hadoop to leverage their large amount of data and reduce cost, increase quality and operational efficiencies.

United Healthcare is mainly focusing on voice based data which is collected from customer desks of call centre. Thus collected voice data is than converted into text data. United also goes for integrated statistical analysis with Hadoop.

D. Big Data at GE

GE, General Electricals, Company helps customers realize greater returns faster from Industrial Internet – a technology framework that connects machines, facilities and people to deliver business value and improve productivity. In a gas/oil industry, with about 1200-turbine wind farm contains approximately 50 sensors with data sampled at different rates [9].

The first level analysis, the real-time analytics within the turbine controller use sensor data collected, and saved in an onboard data historian, every 40 milliseconds to optimize the pitch of the turbine's blades, the conversion of rotational energy into electricity, and to determine whether electricity should be stored in / discharged from batteries or sent to the transmission grid.

The second level of analysis and interaction occurs at the farm. The farm controller receives more than 30 signals from each turbine at 160-millisecond intervals, and its real-time analytics ensure the right combination of turbines deliver predictable power to the utility. In addition, the farm monitoring software processes 200 tags from each turbine at a one-second interval. These real-time analytics evaluate turbine health and performance for the site operations team. Turbine data is then transmitted at one-minute intervals from this and other wind farms to a remote monitoring center miles away.

There, individual turbines and entire wind farms are analyzed to finely tune their process and asset algorithms, and these enhanced analytics are pushed down to the machines. Simultaneously, a high-performance computer cluster is utilizing years of operating data to build predictive models that find correlations and critical issues hidden amongst the tags of thousands of individual turbines. When millisecond granularity is needed to match signatures and patterns found in the streaming asset and process data, queries from the monitoring center can be distributed over thousands of wind turbines' data historians. The insights help operations teams prioritize maintenance, parts warehousing, logistics, and other services for the farm.

Finally, analyses that blend operational data from the farm with financial and other data from the operator's enterprise systems are delivered to the CFO's office in the form of forecasts and power production reports. Similarly, the VP of Operations can access reports on turbine capacity, while field management uses reports that detail a prioritized list of maintenance requirements. All are delivered with a modern user interface on whichever is the most appropriate device for the job.

The key to success for this wind farm lies in the ability to collect and deliver the right data, at the right velocity, and in the right quantities to a wide set of well-orchestrated analytics and provide insights at all levels in the operation. This requires a distributed computing fabric optimized for industrial big data in its many forms and in support of its many different uses. These requirements span the industrial world and are what drive the need for an industrial big data platform.

E. Big Data at Macys.,com

Macys.com is considered the equivalent of a single store at the giant retailer's structure, but it's growing at a 50% annual rate—faster than any other part of the business. The division's management is very oriented to and knowledgeable about IT, data, and analytical decisions. Like other online retailers, Macys.com is heavily focused on customer-oriented analytical applications involving personalization, ad and email targeting, and search engine optimization. Within the Macys.com analytics organization, the "Customer Insights" group addresses these issues, but it also has a "Business Insights" group (focused primarily on supporting and

measuring activity around the marketing calendar) and a "Data Science" organization. The latter addresses more leading-edge quantitative techniques involving data mining, marketing, and experimental design.

Macys.com utilizes a variety of leading-edge technologies for big data, most of which are not used elsewhere within the company. They include open-source tools like Hadoop, R, and Impala, as well as purchased software such as SAS, IBM DB2, Vertica, and Tableau. Analytical initiatives are increasingly a blend of traditional data management and analytics technologies, and emerging big data tools. The analytics group employs a combination of machine learning approaches and traditional hypothesis-based statistics.

F. Big Data at Schneider National

Schneider National, one of North America's largest truckload, logistics and intermodal services providers, has been pursuing various forms of analytical optimization for a couple of decades. What has changed in Schneider's business over the past several years is the availability of low cost sensors for its trucks, trailers and intermodal containers. The sensors monitor location, driving behaviors, fuel levels and whether a trailer/container is loaded or empty. Schneider has been transitioning to a new technology platform over the last five years, but leaders there don't draw a bright line between big data and more traditional data types. However, the quality of the optimized decisions it makes with the sensor data – dispatching of trucks and containers, for example – is improving substantially, and the company's use of prescriptive analytics is changing job roles and relationships.

New sensors are constantly becoming available. For example, fuel-level sensors, which Schneider is beginning to implement, allow better fueling optimization, i.e., identifying the optimal location at which a driver should stop for fuel based on how much is left in the tank, the truck's destination and fuel prices along the way. In the past, drivers have entered the data manually, but sensor data is both more accurate and free of bias.

Safety is a core value at Schneider. Driving sensors are triggering safety discussions between drivers and their leaders. Hard braking in a truck, for example, is captured by sensors and relayed to headquarters. This data is tracked in dashboard-based safety metrics and initiates a review between the driver and his/her leader. Schneider is piloting a process where the sensor data, along with other factors, goes into a model that predicts which drivers may be at greater risk of a safety incident. The use of predictive analytics produces a score that initiates a pre-emptive conversation with the driver and leads to less safety-related incidents.

G. Big Data at Bristol-Myers Squibb

Bristol-Myers Squibb reduced the time it takes to run clinical trial simulations by 98% by extending its internally hosted grid environment into the AWS Cloud. The company has also been able to optimize dosing levels, make drugs safer, and require fewer blood samples from clinical trial patients.

Because clinical trial data is highly sensitive, Bristol-Myers Squibb built a dedicated, encrypted VPN tunnel to the Amazon gateway and configured a virtual private cloud so the environments would be isolated from public customers.

Before moving into the cloud, scientists were using a shared internal environment, so it took 60 hours to run hundreds of jobs. Now that each scientist has a dedicated environment, 2,000 jobs can be processed in 1.2 hours without causing an impact to other members of the team.

As a result of the move, Bristol-Myers Squibb was able to reduce the number of clinical trial subjects in a pediatric study from 60 to 40, while shortening the length of the study by more than a year.

H. Big Data at an International Financial Services Firm

For one multinational financial services institution, cost savings is not only a business goal, it's an executive mandate. The bank is historically known for its experimentation with new technologies, but after the financial crisis, it is focused on building its balance sheet and is a bit more conservative with new technologies. The current strategy is to execute well at lower cost, so the bank's big data plans need to fit into that strategy. The bank has several objectives for big data, but the primary one is to exploit a vast

increase in computing power on dollar-for-dollar basis. The bank bought a Hadoop cluster, with 50 server nodes and 800 processor cores, capable of handling a petabyte of data. IT managers estimate an order of magnitude in savings over a traditional data warehouse. They take existing analytical procedures and converting them into the Hive scripting language to run on the Hadoop cluster.

I. Big Data at Sears

When it comes to the adoption of information technology, Sears was years ahead of most retailers, implementing an enterprise data warehouse in the 1980s while most retailers were still relying on manually-updated spreadsheets to examine their sales numbers. These days the company is using big data technologies to accelerate the integration of petabytes of customer, product, sales, and campaign data in order to understand increase marketing returns and bring customers back into its stores. The retailer uses Hadoop to not only store but process data transformations and integrate heterogeneous data more quickly and efficiently than ever.

The company is now leveraging open source projects Apache Kafka and Storm to enable realtime processing. The main goal of the company is to measure what's just happened.

The company has cited big data's capability to decrease the release of a set of complex marketing campaigns from eight weeks to one week. Faster and more targeted campaigns are just the tip of the iceberg for the retailer, which recently launched a subsidiary, MetaScale, to provide non-retailers with big data services in the cloud.

J. Big Data at The WeatherCompany

A recent partnership between "The Weather Company" and "IBM" will allow companies to better manage the impact of weather on business performance. According to The Weather Company, weather has an economic impact of half a trillion dollars annually in the US alone[2].

The weather data is being collected from more than 100,000 weather sensors and aircraft, as well as millions of smartphones, buildings, and moving vehicles. That data is combined with data from other sources to yield 2.2 billion unique forecast points, and an average of more than 10 billion forecasts on an active weather day. Retailers will be able use the data to adjust staffing and supply chain strategies. Energy companies will be able to improve supply and demand forecasting. Insurance companies will be able to warn policy holders of severe weather conditions, so they can minimize the possibility of car damage in the event of a hail storm, for example.

II. CONCLUSION

It is clear from this research that most of the large organizations from different area are already into big data making full utilization of big data analytics. They are not just using the traditional analytics but also integrated big data separate with them to form a new analytics.

REFERENCES

[1] SAS 2015 Big Data Survey, http://www.sas.com/resources/ whitepaper/wp_58466.pdf
[2] New Vantage Partners, "Big Data Executive Survey: Themes and Trends," 2012.
[3] Thomas H. Davenport , Jill Dyché, International Institute of Analytics 2013 http://www.BigData/ Big-Data- in-Big-Companies.pdf
[4] http://www.informationweek.com
[5] Big Data in Retail Examples by Chunk schaeffer
[6] Mark Van Rijmen, "ups-spends-1-billion-big-data-annually" https"//www.datafloq.com
[7] https://www.mapr.com/customers/unitedhealthcare
[8] http://ebooks.capgemini-consulting.com/dm/Caesars.pdf
[9] http://www.ge.com//Industrial_Big_Data_Platform.pdf