Story of Bigdata and its Applications in Financial Institutions

Phani Bhooshan[1], Dr. C. Umashankar[2]

[1]Research Scholar, Department of Computer Science and Technology,
Sri Krishnadevaraya University, Anantapur, Andhra Pradesh, India
[2]Retired Professor, Department of OR&SQC, Rayalaseema University, Kurnool, Andhra Pradesh, India

ABSTRACT

The importance of BigData is indeed nothing new, but being able to manage data efficiently is just now becoming more attainable. Although data management has evolved considerably since the 1800's, advancements made in recent years that have made the process even more efficient.

Technique of Data mining, is much used in the banking industry, which helps banks compete in the market and provide the right product to the right customer. While collecting and combining different sources of data into a single significant volumetric Golden Source of TRUTH can be achieved by applying the right combination of tools. In this paper Author introduced BIGDATA technologies in brief along with its applications.

KEYWORDS: BigData, Finance, Evolution, Data Science

1. INTRODUCTION

Some Data pundits have mentioned BigData as "large, overpowering, and tremendous amounts of events." In the past, someone said "devastating amounts of information" while studying bubonic plague, which was ravaging Europe. During those days, mathematicians had used stats and were applauded for being the first person to introducing statistical data analysis.

The evolution of BigData includes several preliminary steps for its foundation, and while looking back to 1663 isn't necessary for the progress of data volumes today, the point remains that "BigData" is near term depending on who is discussing it. BigData to Amazon or Google is very different than BigData to a medium-sized healthcare institution, but no less "Big" in the brains of those competing with it.

Foundational steps to the state-of-the-art conception of BigData involve the development of computers, smartphones, WWW, and sensors (IoT) equipment to supply data. Credit providing institutions also played a role, by providing an increasingly large volume of data, and indeed social media changed the way we think of data in nature and still evolving new ways. The progress of modern technology is interwoven with the growth of BigData

1.1. The Foundations of BigData

Day by day increasing Data volume became a problem for some of western governments in the 18th century. Those days statisticians estimated it would take eight years to hold and analyze the data collected during the 18th-century census, and also predicted the data from future census would take more than ten years to understand. In 1881, a young person named Herman Hollerith working with central government bureau, created Machine called Hollerith Tabulating. The invention was using punch cards designed for utilizing the patterns woven by motorized looms. This tabulating mechanism reduced ten years of labor into three months of work.

In 1927, Fritz Pfleumer, an Austrian-German engineer, invented a means of storing information fascinatingly on electromagnetic tape. Pfleumer had developed a method for holding metal strips to cigarettes (to keep smokers' lips from being tarnished by the regular papers accessible at the time), and determined he could use this procedure to create electromagnetic strips, which would then be used to supersede wire recording technology. After experimentations with a variety of materials available during those days, he settled on a delicate paper, barred with iron-oxide powder and coated with lacquer, for his patent in 1928.

During World War II, mid-1940's, Britishers were desperate to decode Nazi signals, which made them invent a machine that scanned for patterns in messages intercepted from the Germans. The device was called Colossus, and scanned 5k

characters per second, reducing the person-hours from several weeks to just in hours. Jokingly Colossus was the first data processing machine. In 1945, John Von Neumann put out a paper on the Electronic Discrete Variable Automatic Computer, in short (EDVAC), the first "documented" discussion on program storage, and laid the foundation of computer architecture today.

It is said these united events prompted the "formal" creation of the United States' NSA (National Security Agency) in 1952, by President Truman, Staff at the NSA were assigned the task of decrypting communication captured during the Cold War. Computers of that time had progressed to the point where they could collect and process data, operate independently and automatically.

2. Big Data and Financial Institutions
Financial institutions are the backbone of any countries economy's health. And Financial institutions are leading the adoption of technology preceded by healthcare and manufacturing companies.

Thus financial institutions have continued to upgrade and patch according to the latest trends and demands from auditors and Controllers. Thus Mainstream Financial institutions enterprise systems became so vast and so fast in terms of depth and breadth are it's unimaginable. Thus creating clusters amongst different systems and repeating and reusing the information according to business demands.

Given the latest technological trends, businesses have determined to combine or merge these legacy systems and eliminate good old systems.

2.1. Challenges with Legacy Systems
Legacy Systems are suitable for the purpose which were they build during their era. However, growing business demand and past pose a risk. Few of the perils and issues listed below.
- Not prepared for Change: Any changes to the existing system would require it to stop and pause to supply patches. This may interrupt business processes and continuity.
- Security of the System: Vulnerability of the system makes it worst target for cybersecurity and data breaches a common issue due to which regulators fine huge sums.
- More demanding Customer: Customers like to use latest gigs and gizmos and may not want to sit on the old swing chair than a bean bag. i.e., customers love to adopt the most recent trends, i.e., buying things on their smartphone than useful old websites with lagging checkouts.
- Cos Effective: Legacy software is outdated, and service providers does not continue to support old version/technologies. Thus leads to high cost in maintenance and effectiveness to the organization.
- Unknown Dependencies: As the people and systems move and SME knowledge go along with these folks, and it would become harder for the organization to switch to new systems.
- Return on Investment: Organization does not like to add more money to the existing old systems as it costs more and returns (RoI) from these systems are very low.

The list goes on, and increase organizations pose more risks with these legacy systems.

3. Organizations Approaches
Once the business has decided to move on with these growing legacy systems, the company could approve one of the below mentioned three approaches:
A. Consolidation and Elimination
B. Consolidation and Remediation
C. Rationalization

Before a company can decide on which approach to adopt, they would have to honestly and accurately understand enterprize level financial and technological targets for their business growth.
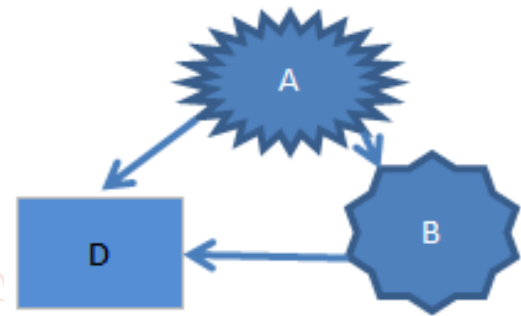

Figure 3.1 Legacy systems and Interlink

Let's take an example with three systems, as shown in the above figure with different technologies exchanging information across. We are not specific to any technologies; it could be data, web or messaging or programmatical. We shall see how each of the approach mentioned would workout in terms of final systems remain in the target operating model.

3.1. Consolidation and Elimination
In this method, first and foremost one has to identify business process and features in each of the system which are classified as to be consolidated. Another critical aspect is existing Data in legacy systems and its structures. Does the old data needs to be migrated as well to the new system.

Once the new system is successfully developed, deployed and tested with necessary signoffs, old data to be migrated as well, after few rounds of production parallel running and gaining user confidence old system would be retired for good.
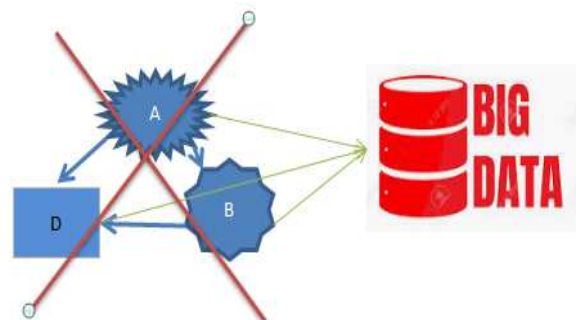

Figure 3.2 Retired Legacy systems and BD

This process might involve SME dedicated input and may cost high at the beginning of adoption. This might work for more significant enterprises with huge budgets and time. However, in the long run, this method would give a perfect Return on Investment(RoI)

3.2. Consolidation and Remediation

This is NOT inline with a similar approach of (i) where identifying common aspects is key to the target model. In this method, we identify loopholes, audit requirements, and regulatory mandates to be adopted for future growth. Here future refers to the long term than the short term.

Rather than trying to apply proprietary patches to the existing system and increasing complexity with more and more features to be added in resolving business needs, audit requirements and more demanding new regulatory requirements to the lagging systems. Some of the functionality and/or specifications may not be possible to be implemented likes of cybersecurity, data secrecy, and modern state-of-art drill up/down facilities.

This leads to the consolidation and remediation process and leads to below target working model fo the sample systems that we had seen in figure 3.1.
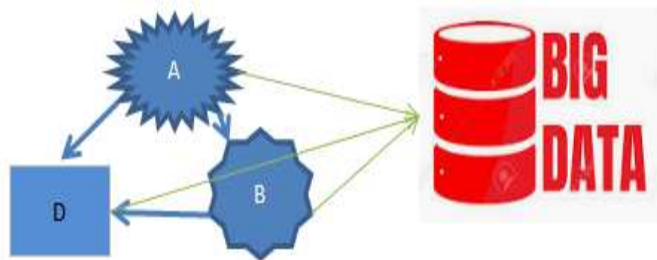

Figure 3.3 Legacy systems and Interlink

In the above final architecture of the systems involved will have one new system which serves new functionalities contemplating existing systems. Down the line, these existing legacy systems may be considered for Elimination as a result of moving all the features to the new system.

3.3. Rationalization

This approach is a hybrid of both 3.1 and 3.2 topics. Where in the best of both the strategies are combined with upgrading existing systems for future.

Step 1 involves finding out all the features of each of the system, along with its corresponding data sets. Next identifying the process to move existing data along with relevant functions to be implemented in the other system with limited or low cost-based approach to add these to other systems.
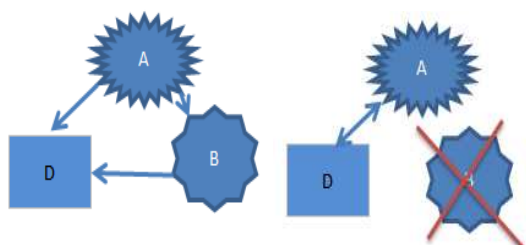

Figure 3.4 Legacy systems Before and After

In the above systems System's B's features are transferred/implemented into combined System A and D. Post-implementation together System A+B+D = A+D nothing less if not more.

This approach may be theoretical, but it's much complicated and depends on the technologies that have been built into the legacy systems.

4. BigData Storage Evolution

Technology does not just mean software which could handle the system. It means a combination of other attributes of the technological ecosystem. One of the main drivers of the latest tech trends is ability to store and retrieve data, fuel to the system. We shall leave early data storage in terms of punch cards, capacitors, and start from Magnetic to Cloud.

Magnetic storage in 1927 Magnetic Stripes were first introduced to store data – called tapes. This was mainstream data storage a decade, though storage capacity has been increased to many folds. In chronological order in which they had invented and increased their capacity were tapes, drums, floppies, and hard disks. Magnetic storage refers to the way how data has been stored. It uses magnetic polarities, to represent a zero or one. Even in the current advanced world some of the organizations still use Tape storage for cold storage for year as part of Regulatory requirement to keep data safe for on-demand basis.

Flash storage is non-volatile memory which increases the performance of the disks to several folds and size has become a tiny of a figure nail to hold few Gigabytes.

Cloud Data Storage is the latest trend and up ticking by most advanced technologies. This storage has become quite popular in recent years due to its high availability and mobility and can be Geographical distributed. Other features of this storage include but not limited to is scalability, low-cost based, energy-efficient.

4.1. Applications of BigData

BigData is transforming the entire financial industry and changing human culture to behavior. As a result of the fast exchange of information and is challenging how people trade, listen to music, and work. The following are some examples of BigData usage:

➢ BigData is being used in healthcare to map disease outbreaks and test alternative treatments.

➢ NASA uses BigData to explore the universe.

➢ The music industry replaces intuition with BigData studies.

➢ Commercial Firms use BigData to study customer behavior and avoid blackouts.

➢ Latest Health Gizmos uses health monitoring wearable's to track clients' activity and provide insight into their health.

➢ BigData is being used by cybersecurity to stop cybercrime.

5. BigData Analytics

Data Visualization techniques are steadily becoming more popular as an essential method for gaining insights from BigData. (Graphics are typical, and animation will become communal. For now, data visualization models are a little inelegant, and the possibility to use some improvements.) Some of the most popular and trending BigData visualization tools are listed below:

➢ Domo

➢ Qlik

➢ Tableau

➢ Sisense

➢ Reltio

6. CONCLUSIONS

To be sure, the Brief History of BigData is not as brief as it seems. Even though the 17th century didn't see anywhere near the Exabyte-level volumes of data that organizations are contending with today, to those early data pioneers, the data volumes certainly seemed daunting at the time. BigData is only going to continue to grow, and with it, new technologies will be developed to collect better, store, and analyze the data as the world of data-driven transformation moves forward at ever-higher speeds.

REFERENCES

[1] Deshpande, M. S. P., and D. V. M. Thakare, 2010. Data mining system and applications: A review. Int. J. Distrib. Parallel Syst., 1: 32-44.

[2] T. AshaLatha, Naga Lakshmi. N, Improving operational efficiencies using Big Data for Financial Services, Volume 18, Issue 4, IOSR Journal of Computer Engineering (IOSR-JCE), Aug. 2016 PP 75-77.

[3] Big Data and Analytics – Wiley Publication, Seema Acharya, Subhashini Chellapan

[4] Thomas H. Davenport Jill Dyché, Big Data in Big Companies, International Institute For Analytics, 2013