Scientific Journal Impact Factor (SJIF) : 1.711
International Journal of Modern Trends in Engineering and Research
ISSN (ONLINE) : 2349 - 9745 ISSN (PRINT) : 2393 - 8161

Big Data and Future of Examination Systems

Tushar Vyas[1], Pankaj Chittora[2], Manish Yadav[3], Shivani Sharma[4]
[1,2,3,4] *Computer Science & Engineering, Vivekananda Institute of Technology-East*

Abstract- With the increase in use of modern tools, the amount of data being generated is also increasing rapidly. The growth of data has affected all fields where data processing is required. This is where big data plays a major role. Big data is a collection of diverse and complex data sets managed in an efficient way using specialized tools. In this paper, we discuss the architecture and challenges in big data analysis for examination system to be used by any state university.
Keywords— big data, data processing, data analysis, challenges, examination system.

I. INTRODUCTION

Today, there is a lot of enthusiasm about the term 'Big Data'. Big Data refers to a group of data that is large and difficult to process using conventional Database Management Tools. Big Data is collected from large no. of sources. Data can be in the form of text, audio, video or picture. Since there is no specific structure of data, we cannot process it using any specific database management system. Thus we require a different approach to process such data i.e. big data, where we use massive parallelism on readily available hardware.

Big Data Analysis is required in almost every aspect of our modern society including mobile services, business, social media, health care etc. Fields like Biotechnology, Astronomy, Geology, etc. are also migrating to big data. In this paper we propose the application of big data in field of education. For instance a traditional examination system of a technical university involving thousands of students. If the data of these students is managed electronically then it would be very difficult to manage and analyze using on-hand database tools.

II. CHARACTERISTICS OF BIG DATA

We can describe big data using the following characteristics:-

1. Volume
2. Variety
3. Velocity
4. Veracity
5. Value

2.1. Volume

The term itself justifies that the content available should be large enough to be termed as big data. In this we talk not only of terabytes but of zetta bytes and bronto bytes. Each and every minute the data is exponentially increasing. Especially if we talk about the social networking, the data sets being developed are too large to be stored and analysed through traditional databases. With big data technology we can now store and use these data sets with the help of distributed systems, where parts of the data is stored in different locations and brought together by software.

2.2. Velocity

This refers to the speed of generation of data (The speed at which the data is processed in order to meet the demands of people). The passing of messages in social media is an example of it. Big data technology allows us now to analyse the data while it is being generated, without ever putting it into databases.

2.3. Variety

The data analysts require the information about the type of data so that effective analysis is possible. Earlier the data was expected to be structured that could fit in the tables and rows schema of conventional databases but now most the available data is unstructured. In the contemporary world maximum fraction of data involves pictures, videos; audios, sensor data etc. are unstructured form of data. It helps the people who are closely analysing the data and are related to it.

2.4. Veracity

Veracity refers to the messiness or trustworthiness of the data. With many forms of big data, quality and accuracy are less controllable (just think of Twitter posts with hash tags, abbreviations, typos and colloquial speech as well as the reliability and accuracy of content) but big data and analytics technology now allows us to work with these type of data. The volumes often make up for the lack of quality or accuracy.

2.5. Value

Then there is another V to take into account when looking at Big Data: Value! It is all well and good having access to big data but unless we can turn it into value it is useless. So you can safely argue that 'value' is the most important V of Big Data. It is important that businesses make a business case for any attempt to collect and leverage big data. It is so easy to fall into the buzz trap and embark on big data initiatives without a clear understanding of costs and benefits.

III. BIG DATA PROCESSING

3.1. Collect

Data is collected from the data sources and distributed across multiple nodes - often a grid, each of which processes a subset of data in parallel.

3.2. Process

The system that uses the same high powered parallelism to perform fast computations against the data on each node. Next, the nodes reduce the resulting data findings into more consumable data sets to be used by either a human being or machine.

3.3. Manage

Often the bigdata being processed is heterogeneous, originated from different transactional systems .Nearly all of that data needs to be understood, defined, annotated, cleansed and audited for security purposes.

3.4. Measure

Companies will often measure the rate at which data can be integrated with other customer behaviours or records, and whether the rate of integration or correction is increasing overtime. Business requirements should determine the type of measurement and on- going tracking.

3.5. Consume

The resulting use of data should fit in with the original requirements for the processing. For instance, if bringing in a few hundred terabytes of social media interaction demonstrates whether and how social media data delivers additional product purchases, and then there should be rules for how social media is accessed and updated. This is equally important for machine-to-machine data accesses.

3.6. Store

As the "Data-as-a-service" trend take shape, increasingly the data stays in a simple location, while the programs that accesses it move around, whether the data is stored for short-term batch processing or longer–term retention, storage solution should be deliberately addressed.

3.7. Govern

Data governance encompasses the policies and oversight of data from a business perspective .As defined, data governance applies to each of six preceding stages of big data delivery.

IV. CHALLENGES IN BIG DATA

4.1. Heterogeneity and incompleteness

Machine analysis algorithms expect homogeneous data and cannot understand nuance. Even after data cleaning and error correction, some incompleteness and errors in data are likely to remain.

4.2. Scale

Managing large and rapidly increasing volumes of data has been a big challenge for the past few decades.

4.3. Timeliness

The larger the data set, the longer it will take to analyze.

4.4. Privacy

Managing privacy is effectively both a technical & a sociological problem that must be addressed jointly form both perspectives to realize the promise of big data.

4.5. Human Collaboration

The data system needs to be designed such that it accepts the distributed expert input & support their collaboration.

IV. BIG DATA IN EXAMINATION SYSTEM

The modern day exam system is comprised of mid-terms, class tests and end term exams. Each has a different method of evaluation and may also have different structures. Recently the govt. focused on giving maximum priority to percolate education for all the citizens of our country. The main challenge in fulfilling this target is the proper distribution of education resources ad well as monitoring and administering the whole system.

A part of this problem can be solved by establishing a framework on which all exam systems will be administered in a structured manner. This framework will use the concept of big data. This will allow better analysis of large volumes of data gathered from various exam centers, improving the efficiency of the overall education system.

According to our proposed architecture exams will be conducted online and the data i.e. solutions will be evaluated on a central server and results would be made available in shorter time periods. Assuming a large no. of students appearing for different exams, we would definitely need big data analysis. The Steps that can be involved in processing of such data is as follows:

1.) *Collect*- Answers to different questions in an exam will be attached to student profiles and will be saved on the central server in encrypted form.

2.) *Process*- Once data from all the exam centers is collected, the central server then processes it and evaluates results for each student.

3.) *Manage*- The evaluated results are then declared on a single server which is publically accessible.
4.) *Measure*- As the amount of data to be integrated is very large, certain data load measures must be adopted.
5.) *Consume*- The massively generated data must fit within the original requirement for the consumption of the data.
6.) *Store*- The final data generated will be stored for a time specified by the different examination centers.
7.) *Govern*- The finalized data should be governed with the policies and regulation specified according to the business purpose.

V. CONCLUSION

As the importance of big data has been realized in various aspects, one can notice how it can be specially applied on our examination systems. By using big data in the existing system we can further refine it and make it more efficient. Some of the current loopholes can be rendered by this proposal. The scenario is clear with the application prospective. The power to store, organize and manipulate data exists only with the Big Data Systems.

REFERENCES

[1] Mahesh G Huddar, Manjula M Ramannavar, "A survey on Big Data Analytical Tools", IJLTET
[2] P. Zikipoulos, T.Deutsch, D. Deroos, "Harness the power of Big Data", 2012
[3] Priyanka Srivastava, "Big Data and Data Science : Case Studies",IJRSI, Volume I, Issue II, July 2014
[4] M. Jensen, J. Schwenk, N. Gruschka, and L. Lo Iacono, *On Technical Security Issues in Cloud Computing*. IEEE, 2009.
[5] Greg Boss, Padma Malladi, Denis Quan, Linda Legregni, Harold Hall, "Cloud Computing", http://www.ibm.com/developerswork/websphere/zones/hipods/library.html, October 2007
[6] G. Frankova, *Service Level Agreements: Web Services and Security*, ser. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, vol. 4607.
[7] "Service Level Agreement and Master Service Agreement", http://www.softlayer.com/sla.html, accessed on April 05, 2009.
[8] M.V. Luis, R. M. Luis, C. Juan, L. Maik. A Break in the Clouds: Towards a Cloud Definition. Computer Communication Review, vol.39, pp.50-55, 2009
[9] http://hadoop.apache.org.

