IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 7, September 2014.

www.ijiset.com

ISSN 2348 – 7968

An analysis of Big Data usage in Big Organization

C.Vidhya[1]

[1] Department of Computer Science, St.Joseph's College/Bharathidhasan University,
Trichy, Tamilnadu, India

Abstract

Big Data consist of billions and trillions of records from different servers which may be from Business, Education, Service, Social network and mobile data. Each organization is maintaining big data server to manage their Big Data in very efficient manner. They are in need of some consultancy which will be the total responsible for managing the organization's Big Data. Applications for Big Data servers are still developing themselves to attract the Big Organization by their Scalability level and Storing Techniques.

Keywords: *Big Data, Volume, Variety, Access, Storage, Map reducing.*
Submission.

1. Introduction

Big Data is "Massive Volume" Of data. Big Data is very large and it may be structured, semi-structured, unstructured so it very difficult to process. Big Data has been used to convey all sorts of concepts like Large amount of data, Social media analytics and real time data. The characteristics of big data are Amount of data(Volume), User access(Communication), Fast retrieving(Speed of generating), Analyzing data(Meaningful analysis).

Big Data Management is all about maintaining the data correctly and making to work it accurately for a particular organization through a analytic agency. It provides some set of procedures and tools that consistently defines and manage the data according to a particular field.

2. Observations

It has been analyzed that world's all popular social sites are relied with Big Data only and it is successfully undergone with the help of business consultancy. Each and every business consultancy is providing valuable solution to the organization and managing their data accurately, not only with the help of human sources and also with the help of some software tools. Big data management encompasses people, process and technologies that ensure accuracy, uniformity and uniqueness of data entities shared by multiple business units within an organization.

Analytics of Big Data involves collecting those data in order to create meaningful information. All this data will be in various shapes, sizes and forms, represents valuable in formations that needs to be standardized and stored in a meaningful way. Analytics are to provide some set of procedures and tools that consistently defines and helps in managing data correctly.

- "ComScore" is a one of the big business consultancy.
- "Hadoop" is act as a server for Big Data management.

3. Experiment Analysis

This analysis is explaining 2main source involved in Big Data management.

a) ComScore:

Normally, world's big organization cannot directly approach the big data, they must have some business analytics, which will provide services to the organization (or) enterprises.

ComScore is generally defined as "a business analytics which provides business solution and data marketing" for world's largest enterprises, agencies and publisher. ComScore is an "American internet analytics company" providing marketing data and analytics to many of the world's largest enterprises and agencies.

ISSN 2348 – 7968

ComScore is a global leader in measuring digital world because it is playing the main role of storing and managing large amount digital data. It is upfront about collecting user data (or) the software ability to track all of the user's internet traffic including normally secure connection used to communicate banking and other confidential information..

ComScore has Billions of users as it is maintaining very accurate data for the management. **Twitter** is a very important client for ComScore.

- Maintaining Digital data for Twitter(Social network).
- Digital data(Photos, Videos etc).
- Maintaining 1TB(terabyte) data for every 1hour.

b) Hadoop

Also called as Apache hadoop. Hadoop act as a server to manage all the big data through application. It is a framework for running application on large cluster built of commodity hardware and it is providing application with "Reliability" and "Data motion". Hadoop clusters make distributed computing readily accessible to the java community

Hadoop provide 2 important paradigm in creating application called,

- Map reducing technique.
- Distributed file system(DFS).

Map reducing technique:

In this technique, here the application is divided into many small fragments of work and each will be executed (or) re-executed on any node in the cluster. It allows the parallel programming model to be used quickly and efficiently. The map/reducing parallel programming model is the simplest of all parallel programming model, which is very easy to learn than message passing.

The input to mapper is a file and the output is another file. The input to reducer is the set of mapper output files. And moreover hadoop provides a java API for executing map/reduce programs. Map reduce can use data from central storage file system.

Distributed File System(DFS):

Distributed file system is to store data on the computer nodes and providing very high Bandwidth across the cluster. Both map reduce and hadoop DFS are designed to avoid failure and automatically handle the framework.

Client: Linkd in

Linkd in is one of the popular social network connecting people.

4. Literature Review

Big Data Analysis:

Big data analysis is analyzing the big data from its all view and making use of it in very efficient way. Analysis of big data involves many steps to get better result. Big data still needs some take challenge to fulfill the analyzing part [5].

LLGrid Map Reduce:

LLGrid map reduce will enable the data to map or reduce for any type of language using a simple command. This type of map reduce can use data and other information from central storage file system [1].

Master Data Management(MDM):

MDM generally defines a kind of method which gives meaningful context to big data and to view a critical data entities such as product, customer, employee, location etc [7].

Recommendations: Cultivating Big data Adoption:

To extract more value from big data, it offer a broad set of recommendations to organizations as they proceed down the path of big data. Mass digitization, one of the forces that helped to create surge in big data [8].

Revolution analytics and hadoop:

As the amount of data- especially unstructured data- collected by organizations and enterprises explodes, Hadoop is emerging rapidly as one of the primary options for storing and performing operations on that data [6].

ISSN 2348 – 7968

V.Future Enhancement:

In future it would be good if there is enhanced to increase the speed of storing and retrieving the data. These enhancement will increase the no of clients for business consultancy with good name. And also to improve the security to big data for large organization, this will improve the quality of analytiocs.

VI.Conclusion:

Thus the analysis has successfully done with researching the importance of Big Data usage in Business analytics for big organization. Increasing the number of users will improve the quality of business consultancy and their standard.

Refernces:

1. Chansup Byun, William Archand, David Bestor, Bill Beregeron, Mathew Hubbell, Jeremy Kepner, Andrew McCabe, Peter Michaleas, Julie Mullen., Driving Big Data With Big Compute.., MIT Lincoln Laboratory, Lexinton, MA, U.S.A.
2. Alan Misolve, Massimiliano Macron, Krishna P. Gummadi, Peter Druschel, Bobby Bhattacharjee., Measurement and Analysis of Online Social Networks, University of Maryland, Saarbrucken 66123, German.
3. Juha K. Laurila, Daniel Gatica-Perez, Imad Aad, Jan Blom, Olivier Bornet. Trinh-Minh-Tri do, Olivier Dousse, Julien Eberle, Markus Mirttinen., The Mobile Data Challenge: Big Data for Mobile Computing Research., Nokia Research Center, Lausanne, Switzerland.
4. www.confhub.com
5. Challenged and opportunities with Big data., A white paper developed by researchers from united states.
6. Advanced ‘Big Data’ Analytics with R and Hadoop, A white paper from Revolution Analytics.
7. Mastering Big Data: The next big leap for master data management, A white paper from Coginzant.
8. Gauvrav vaswani, Student, VESIT, Mumbai, India, Amradha Bhatia, Faculty, Computer tech dept, Mumbai, India., A real time approach with Big data.
9. Gilbert wordracek Thorsten holz, technical university Vienna, Austria., Engin Kirda, Institute Eurecom., Sophia antipolis., A practical attack to de-anonymize social network users.
10. Yong yeal ahn,Hawoong jeong, department of physics, KAIST, deajeon, korea., Seung yeop han, division of computer science, KAIST deojeon korea., Haewon kwak, Sue Moon, division of computer science, KAIST deojeon korea., Analysis of topological characterstics of huge online social networking services.
11. Alan Mislove, massimiliano macron, Krishna P. Gummadi, peter druschel, Bobby Battacharjee, MPI for software systems, campus E14, Saarbrucken germany., Measurement and analysis of online social networks.
12. Big Data meet Big data analytics, A white paper from SAS.
13. Big Data: Beyond the hype, A white paper from Datastax.

First Author Pursuing master of philosophy in St.joseph’s college and now working with N.S.College, Theni. Research area is Distributed Database Management. Researching the storage part of Big Data Under the guidance of Dr.C. Balakrishnan.