Review Paper on Big Data

Swapnil A. Wankhade
Department of computer science & Engineering, Sipna college of Enggi. & Tech. Amravati, India.

Abstract – Big data is a large data storage capacity in this day it will be storing data ,distibute , capturing the picture and exchanging data in wast manner they will be access.Big data day by day it will it more popular and more new technique will be adopted. They are accessible on internet, social networking site to it can store large amount of data in peta byte and exabyte. Big data often refers to simply handle of predictive analytics and user behavior analytics. they more higher data analytics procedure that extract value form data. Big data can be include V3 process called velocity, volume, variety . The big data they will be use in banking , companies and education on open source this paper aim to aware of big data analyze and applying different tools to it handle.

Index Terms – Big data, Analyzing.

1. INTRODUCTION

In this situation almost every where or time they will use 'Big data'. After work will done data is lost. The organizationlost ability and storage information and knowledge. Big data are storage data capacity increase and method of data collection. Large amount of data have easily available on internet. Every second create data and it will be store and analyzed and it become easier to store the data in big data. Big data is a process to handle to data set volume, variety, velocity to difficult to handle captured, process and analyzed conventinol method tools and tech. which can be in big data. The big data are more correctness is leadto confident to take decision and big data reduce the cost and also they will beremove the risk in the data of big data. Analysis of big data they will be find correlation to prevent the harmfull process. every day 2.5 Exabyte data will be create at time or day. So this way it will be acces and work will be done in "big data".

2. DEFINITION : ON BIG DATA

The big data is data sets that are so big or complex that traditionaldata processing application software is lack of competness to deal them. Big data is process that will be access or refer to data sets or multiple of data sets these size beyond the software tools capature, managed, within a time. Big data size are permanent moving the target . Big data is a set of techniques and technologies that create a new forms to integrate to uncover large hidden values from large datasets that are diverse,complex and of the massive scale. Big data most analysts and practitioners current refer data sets from 25-55 terabytes(10 12 or 1000 gigabytes per terabyte) to multiple petabytes(1015 or 1000 terabyte per petabyte)are the big data.In the 2001 research report and releted lecture ,META Group(now the Gartner) analyst "DOUG LANEY"[4] defined growth challenges and opportunities as being three dimentsional means to incresase the volume(amount of data), Velocity(speed of data in & out), and last one is variety(range of data type and source).

3. CHARACTERISTICS : ON BIG DATA

The big data whose scale, separation, root change and timeline require to use new technical architectures, analytics and tools in order to insight that free sources of big data .The main characteristics of big data are given below That's also called as 'V3'. Detail description of 'v3' that can be follow.

Volume : The avability of data is created is very important in this context. It is the size Of data the which analysis the value and ability of data under consideration which actually be considerd big data or not. The big data related to size and his char. The data will store megabytes gigabytes to petabytes.

Variety : The next step of the big data is variety. this means different type of data and sources of data. Category to which big data belongs to is essential fact that need to known by data analysts. Big data draws from text, images, videoalso it completes missing pieces through data fusion.

Velocity : The Velocity is process to check to speed of generation of data or how Fast data generated and process to complete the demands and challenges to Which acess in path in the growth and development.

4. BIG DATA ANALYZING PROCESS

After the big data storing process then it come anlayis process there are many critical requirements for big data. The first one is fast data loading and accesing data. The big data and network traffic interferes with the query exeute during data will be loading it is compulsory to deduct the loading time. And the next requirement is fast query process. In this process fulfillthe requirement of heavy workload and realtime request so many queries areresponse time are dangers. Thus data placement structure must capable for handling high qurery processing speed as the amount of queries fastly increase The third next requirements for big data processing is highly proper management Of storage space. They are fastly growth in user activity in demand for storage Capacity and power, data storage well management during process issue how To store data that space utilization is maximized be addressed. The final requirement is strong adaptivity to highly big workload patters. The big data sets are analyzed by different user and application

for different purpose and in various way. The map reduce is a parallel process program model and develop by a map & reduce of functional languages which is accessible for big data processing. It is core of Hadoop and perform the data processing and analysis function. MapReduce process is work on adding more computer and handler also it increase the power efficiency and large space for in a specific computer. in another way to express it is divide the operation in part of task in stages and stage execute in parallel in process to reduce the time.

5. LITERATURE REVIEW ON BIG DATA

S. V. Phaneendra & E. M. Reddy : "Big Data- solutions for RDBMS problems. Apr 23 2013[1]. survey" The few years ago of day the Data will small and storage capacity that will be handle by the RDBMS. But in this time data are so large and they are difficult to handle large data Through the Rdbms method or tools. So this time use 'Big Data'. he say that big data differently change from other data such as Volume, Velocity, Variety, etc. Hadoop process large number of data sets. They will focused on challenges and need to handle the big data.

K.K.Reddi & D. Indra :"Different Technique to Transfer Big Data : survey".Aug.2013 [2]. They also focused on the knowledge that big data Is collection of structured, semi-structured, non structured homogenous And hetrogenous data. the author say that to access nice model to handle Transfer of large amount of data in networking. these transfer is permited to idle bandwidth available. The good process model use to store and transfer to access staging server. They also advice that new algo. are require to sending big data and solving problem means security.

Jeffrey Dean : MapReduce on simplified data.Jan 2008 [3] The MapReduce are perform on large cluster of commodity machine and is big scalable In a perfect manner MapReduce combinationprocess many terabyte of data thousand machine. The programmer and system easy to handle. Many mapreduce program have be construct and upwords thousand mapreduce job are execute on google cluster each day.Author proposes proper utilize data processing on large Clusters.

J. S. Ward : " The big data storing process Sept 2013 [5] Did survey of big data definition. The big data has two idea are "Data storage and Data analysis". This create question how the big data is Different from other data process technique. That the answer are in this process they are big implies significance complexity and challenges. The lack of consistent definition to meet ambiguity and hampers discource related to big data.

6. CONCLUSION

In the review paper on big data they are completely different topic. The paper decribes the cocept of big data along with V3 Volume, velocity and Variety of big data. This paper also proper study on big data processing problems. The Big data are challenges for better work and fast working process of big data. So they more data will store big data in that manner they will be easy to access. The big data analysis can be apply to converting business change and large Decision making by using advance analytic processes on big data, valuable knowledge. So the big data is large data application process. Thus we have focused on big data.

REFERENCES

[1] Google wikipidia on big data information and analyzing data www.google.com

[2] "DOUG LANEY": MATA DATA[4]

[3] S. V. Phaneendra & E. M. Reddy : "Big Data- solutions for RDBMS problems. Apr 19 23 2013[1].

[4] K.K.Reddi & D. Indra :"Different Technique to Transfer Big Data : survey".Aug.2013 [2].

[5] J. S. Ward : " The big data storing process Sept 2013 [5]

[6] Jeffrey Dean : MapReduce on simplified data.Jan 2008 [3]