

Handling Database through BIG Data Technology

Er. ManojSaraswat¹Er.A.K. Khan²

^{1,2}Asst.Professor

^{1,2}S S College of Engineering, Udaipur

Abstract—Handling of Big Database is need of many organizations, companies, and researchers to deal with big data volumes efficiently. Examples include web maintenance, specific scientific programs, and social networks (such as Facebook, Amazon etc). A new and enraging technology for processing of data in search enginisis big data-technology name is Hadoop Map-Reduce. At the starting Hadoop was not accepted in industry for basic lacunas in it . But now it is past. Now many methods are used with Hadoop- MapReduce technology to improve efficiency in Databases, First, at the starting use and need of it is specified for large data processing. Then, we will categories on different data management methods,

Keywords:Hadoop, Map-Reduce, BIG DATA.

I. INTRODUCTION

“Big Data” is data whose scale, diversity, complexity requires new architecture, techniques, algorithms, and analytics to manage it and extract value and hidden knowledge from it.

Data Volume: It shows the requirements of arranging data in coming years
44x - increase from 2009 -2020

From 0.8 Zetta bytes to 35 Zb

Database is increasing exponentially

Let us take an example to give insight of a Big Data. We take a case of Bike verses a cab. Bike provides a means of transportation but we have to take care personnel handling and self-control also. On the other hand car is uses publicly .It also provides the transportation with full facility[1].

So cloud is moreover like a Car (we use as per our requirement and as per our demand).

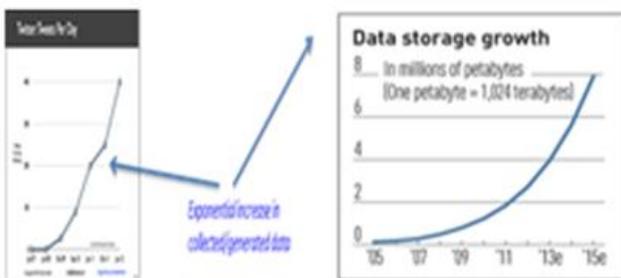
Now we move toward the details of Big Data computing[1].

Characteristics of Big Data:

A. Volume :

It is the most commonly and widely used form of Data. In this the working members of all the applications and functions from the web browser. A short or no code is required for the working of these application.

This shows the exponential growth of data handling in today’s world.



B. Complexity (Variation):

In it the servers or run time environments are provided on demand & developed for specific customer applications.

- (a) Various formats, types, and structures
- (b) Text, numerical, images, audio, video, sequences, time series, social media data, multi-dim arrays.
- (c) Static data vs. streaming data

A single application can be generating/collecting many types of data

Well known examples are goggle search engine.

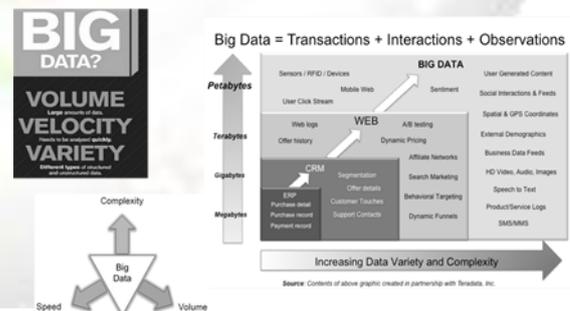
C. Speed (Velocity): In this the

- Data is begin generated fast and need to be processed fast
- Online Data Analytics
- If responded late, opportunity is missed.
- Bye it in less time most data’s are searched and linked together and seen on single platform[2].



Many big corporate houses (like Amazon, IBM, Facebook) are working for using this technology on working of Big Data.

Big Data: 3V’s



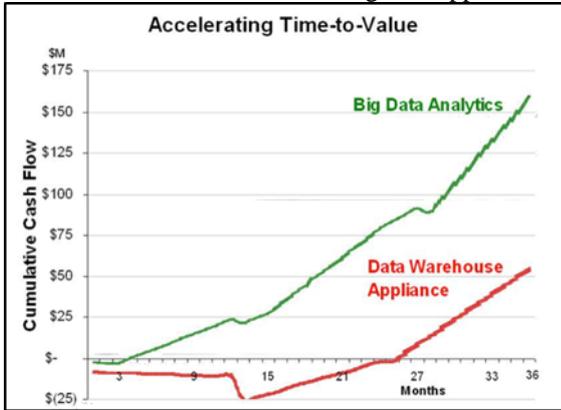
The new implementation of one more characteristic is Veracity- to remove duplication in data if found in the data-base available for work and also remove inconsistency and latency in processing of them.

II. VALUE OF BIG DATA ANALYTICS

The key benefits of using BIG Data are that it provides the

following advantages:

- (a).Big data is more real-time in nature than traditional Data Ware applications.
- (b).Traditional DW architectures (e.g. Exadata, Teradata) are not well-suited for big data apps[4].
- (c)Shared nothing, massively parallel processing, scale out architectures are well-suited for big data applications.



This analysis showing the value of money relating with use of technology, the use of and comparison of BIG Data and traditional databases.

III. HADOOP MAPREDUCE

The technology use in it to implement in database and the parameters used are (a,m).

These are map reduce parameters.

The file system used in it are HDFS(Hadoop Distributed File System)[3].

The reduce function produces zero or more aggregated results.

REFERENCES

- [1] Okcan and M. Riedewald. Processing Theta- Joins using MapReduce. In SIGMOD.
- [2] Thusoo e t al. Data Warehousing and Analytics Infrastructure at Facebook.
- [3] Hadoop,<http://hadoop.apache.org/mapreduce>
- [4] S.Wu, F. Li, S. Mehrotra, and B. C. Ooi. Query Optimization for Massively Parallel Data Processing.