BIG-DATA

Kanimozhi.S (CSE department) kowshalya.M (CSE department)

Contact no: 9786053950

E-mail id:kanimohzi738@ gmail.com

ABSTRACT:

INTRODUCTION:

Big Data is a broad term for data sets so large or complex that traditional data processing applications are inadequate. Challenges include analysis, capture, data duration, search, sharing, storage, transfer, Visualization, and information privacy.

Accuracy in big data may lead to more confident decision making and better decisions can mean greater operational efficiency, cost reductions and reduced risk.

Characteristics:

Volume-The quantity of Data that is generated is very important in this context. It is in the size of data which determines the value and potential of the data under consideration.

Variety-This means that the category to which Big Data belongs to is also a very essential fact that needs to be known by the data analysts. This helps the people, who are closely analyzing the data.

Velocity-It refers to the speed of generation of data or how fast the data is generated and processed to meet the demands and the challenges which the ahead in the path of growth and development.

Complexity-Data management can become a very complex process, These large volume of data need to be linked, connected and correlated in

order to grasp the information that is supported to be conveyed by these data.

Architecture:

In 2004, google published a paper on a process called Map Reduce framework provides a parallel processing model and associated implementation to process huge amounts of data. The framework was very successful, so others wanted to replicate the algorithm. The frame work was adopted by an apache open source project named Hadoop.

Conclusion:

The availability of big data, low-cost commodity hardware, new information management and analytic software produced a unique moment in the history of data analysis. It gains in terms of efficiency, productivity, revenue and profitability.