Abstract

Organizations generate large amount of data each day which involves storage of data, processing of data and retrieval of data for other purpose and usage. But an individual organization or an enterprise finds this difficult i.e. they cannot handle this data and thus the useful data remains useless for a longer duration resulting in waste of storage area. Thus I propose an approach for processing and handling the data from various sources in an efficient manner. This approach can be used for Business Intelligence where the data can be processed and can provide ideas about the popularity, cost and feedback about the product released by the enterprise. The Hadoop technology is used for this purpose.
References

- "MapReduce: Simplified Data Processing on Large Clusters"; Jeffrey Dean and Sanjay Ghemawat.
- "A Performance Analysis of MapReduce Task with Large Number of Files Dataset in Big Data Using Hadoop"; Amrit pal, Kunal Jain, oinkiAgrawal, Sanjay Agrawal.
- Big Data for Business Managers - Bridging the gap between Potential and Value, AnmolRajpurohit, Department of Computer Science, The LNM Institute of Information Technology, Jaipur, India, anmol@lnmiit.ac.in.
- "Reducing the Search Space for Big Data Mining for Interesting Patterns from Uncertain Data"; Carson Kai-Sang Leung, Richard Kyle MacKinnon, Fan Jiang, Department of Computer Science, University of Manitoba, Winnipeg, MB, Canada, kleung@cs.umanitoba.ca.
- "Store, Schedule and Switch – A New Data Delivery Model in the Big Data Era"; Weiqiang Sun, Fengqin Li, Wei Guo, Yaohui Jin and Weisheng Hu, State Key Laboratory of Advanced Optical Communications Systems and Networks Shanghai Jiao Tong University, Shanghai, 200240, China, sunwq@sjtu.edu.cn.

Index Terms

Computer Science
Information Science

Keywords

Big-data Hadoop Mapreduce Hadoop Distributed Filesystem Business Intelligence.