Abstract

The rapid growth in the field of technology in the last 10 years has generated a large amount of data which is structured, semi-structured, as well as unstructured in nature. This large amount of data has been generated by the excessive use of the Internet, the growing popularity of the social networking sites like Facebook, Twitter, Quora, Linkedin etc and there are many more reasons behind this abrupt growth of data. The emerging trend of E-commerce in last few years has fastened the rate of growth of the data. The data has been growing rapidly in Volume, Variety, Velocity and Veracity. This emerging trend coined the term Big Data. The Big Data is often unstructured and real-time analysis is required to analyze and process the data. This evaluation calls for the new system architecture for the storage, transmission, analysis and processing of the data. In this paper we present a literature survey of the Big Data and Big Data analytics. The paper is divided into various sections which comprises of the definition of Big Data and its growing importance, Big Data challenges and Big Data Analytics. We also present a detailed survey of analytics platform like Apache Hadoop. Finally we delineate the evaluation and research direction for the Big Data System.
References

9. “Real-Time Stream Processing as Game Changer in a Big Data World with Hadoop and Data Warehouse”.
Index Terms

Computer Science  Databases

Keywords

Big Data, Big Data Analytics, stream processing, batch processing, Apache Hadoop.