Abstract

Big data refers to a large quantity of data that has to be processed at one time. With the advancement of social media and the virtual world, a vast amount of data is created every second. A technique has to be designed to effectively process this ever-increasing collection of information. One such algorithm is the MapReduce algorithm. The result or output of the algorithm provides useful insights about the data used as input and can be further used for Decision Making and Prediction algorithms. Also, new data is generated frequently for Big Data processing. Hence, MapReduce implementations must not only be accurate but also as instantaneous as possible. This paper discusses not only the details of Map Reduce Algorithm; it also suggests architecture called Dedicated Client Architecture which would increase the efficiency of the algorithm.

References

Dedicated Client Architecture in MapReduce and its Implications on Performance Considerations

- Big Data, Bigger Opportunities, Author: Jean Yan, April 2013.

Index Terms

Computer Science

Data Mining

Keywords

Big Data  Map Reduce  Client Server Architecture  Dedicated Client