Abstract

The generation of technology and requirement fulfill the demand of digital universe data. Day to day the digital universe data are exploded in terms of megabyte and petabyte. The exploding rate of data demands the new generation of technology such as big data processing. In this paper optimized the performance of map reduce programming model for the enhancement of data processing. The modified model of programming used clustering technique. the clustering technique incorporate the process of map data in terms of task group. The task group of map data correlated with different index of data for the processing of data node. The proposed model implemented in Hadoop framework and programmed in java. For the evaluation of performance used three standard datasets and measure the processing time and count value of file.

References


Index Terms

Computer Science  Distributed Systems

Keywords

Big Data, Hadoop, MapReduce, Clustering, Optimization