Algorithms for Intelligent Systems
Series Editors: Jagdish Chand Bansal · Kusum Deep · Atulya K. Nagar

Dinesh Goyal · Valentina Emilia Bălaş ·
Abhishek Mukherjee ·
Victor Hugo C. de Albuquerque ·
Amit Kumar Gupta *Editors*

Information Management and Machine Intelligence

Proceedings of ICIMMI 2019

Springer

Contents

Performance Improvement of Heterogeneous Cluster of Big Data Using Query Optimization and MapReduce

Pankaj Dadheech, Dinesh Goyal, Sumit Srivastava, Ankit Kumar, and Manish Bhardwaj

1 Introduction

Hadoop is scalable and capable of managing substantial data volumes with same or homogeneous kind of clusters since the data movement and processing capacities of servers remain same all around the network. In case of heterogeneous clusters, each node comprises of a host with different speeds of storage and processing capabilities. The task of processing tasks then is changed from slow-end nodes into high-end nodes. This strategy works good once the number of information to be processed is less or job load is low, but it fails to improve the rate of job processing on Hadoop heterogeneous clusters if processing and data involves huge volumes of data sets. The improvement of the hardware product is another issue where processors can be improved but it reveals quite expensive. Another solution for this matter is to improve performance of those heterogeneous clusters using different analytical techniques to analyze structured and unstructured information [1]. The MapReduce algorithm

P. Dadheech (✉) · A. Kumar
Department of Computer Science & Engineering, Swami Keshvanand Institute of Technology, Management & Gramothan, Jaipur, Rajasthan, India
e-mail: pankajdadheech777@gmail.com

A. Kumar
e-mail: iiita.ankit@gmail.com

D. Goyal · M. Bhardwaj
Poornima Institute of Engineering & Technology, Jaipur, Rajasthan, India
e-mail: dinesh8dg@gmail.com

M. Bhardwaj
e-mail: manishbhardwaj.it@gmail.com

S. Srivastava
Manipal University Jaipur, Jaipur, Rajasthan, India
e-mail: sumit.310879@gmail.com

D. Goyal et al. (eds.), *Information Management and Machine Intelligence*,
Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-4936-6_9