



YARN Schedulers for Hadoop MapReduce Jobs: Design Goals, Issues and Taxonomy

Buy Article:

\$68.00 + tax

(Refund Policy)

ADD TO CART

BUY NOW

Authors: Kotikam, Gnanendra; Selvaraj, Lokesh

Source: Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science), Volume 16, Number 6, 2023, pp. 44-55(12)

Publisher: Bentham Science Publishers

DOI: <https://doi.org/10.2174/2666255816666220831125012>

[Abstract](#)
[References](#)
[Citations](#)
[Supplementary Data](#)

Objective: Big Data processing is a demanding task, and several big data processing frameworks have emerged in recent decades. The performance of these frameworks is greatly dependent on resource management models.

Methods: YARN is one of such models which acts as a resource management layer and provides computational resources for execution engines (Spark, MapReduce, storm, etc.) through its schedulers. The most important aspect of resource management is job scheduling.

Results: In this paper, we first present the design goal of YARN real-life schedulers (FIFO, Capacity, and Fair) for the MapReduce engine. Later, we discuss the scheduling issues of the Hadoop MapReduce cluster.

Conclusion: Many efforts have been carried out in the literature to address issues of data locality, heterogeneity, straggling, skew mitigation, stragglers and fairness in Hadoop MapReduce scheduling. Lastly, we present the taxonomy of different scheduling algorithms available in the literature based on some factors like environment, scope, approach, objective and addressed issues.

Keywords: Hadoop map reduce; YARN schedulers; energy consumption; fair scheduling; scheduling issues; virtualization

Document Type: Miscellaneous

Publication date: July 1, 2023

[More about this publication?](#)