

ملخص البحث باللغة الإنجليزية:

Hadoop is the most economical and low cost software framework for storing and processing large data sets in a distributed manner. Using MapReduce, the Hadoop distributed file system (HDFS) enables for distributed storage and parallel processing of large data sets. Hadoop's current implementation, on the other hand, assumes that computational nodes in a cluster are distributed evenly. The major causes of stragglers in heterogeneous Hadoop clusters are load imbalance during storage, resource contention during task scheduling, hardware degradation due to its excessive use, and software misconfiguration during cluster management. Hadoop's performance suffers in a heterogeneous environment due to hardware heterogeneity. The publication summarizes and evaluates major papers in our research. It introduces a new classification taxonomy that divides current straggler management and mitigation solutions into two categories. It examines and analyzes state-of-the-art studies and, based on their findings, determines their limitations. Finally, the systematic

review examines outstanding concerns and future research paths for managing and mitigating struggles by artificial neural network

وتفضلوا سيادتكم بقبول فائق الاحترام والتقدير،،

عميد الكلية ،،

أ.د/ محمد حلمي عبدالعزيز خفاجي