

Abstract

Big Data or Hadoop system is commonly used by big company such as Google, Yahoo, etc. Hadoop itself is a software framework with a Java and open-source basis for processing massive distributed data and run in a cluster which is contain with distributed computer. Hadoop architecture had two layers MapReduce and Hadoop Distributed File System (HDFS). MapReduce is a framework from distributed application while Hadoop Distributed File System is a distributed data. *Fair Scheduler* is one of the job scheduler being used in Hadoop system and has a multi-node system which is contain two servers *main server* and *slave server* then, Fair Scheduling had a characteristic dividing the resources evenly to job which is given by the user. To raise the scheduling performance ini parameter *Task fail rate*, *Job Throughput*, and *Average Completion Time*. *Classification and Optimization based Scheduler for Heterogenous Hadoop Systems* (COSHH) is a scheduler with the purpose to raising the performance from Fair Scheduling with the same logic but, focus on heterogeneity aspect from job sheduler. And hoping that will raising the performance especially in *Average Completion Time* parameter. *Classification and Optimization based Scheduler for Heterogenous Hadoop Systems* (COSHH) had the best performance compare to *Fair Scheduling* in executing the *heterogenous* jobs. And had the best *Average Compelction Time*. With the lowest number in 49,4 Second compare to *Fair Scheduling* in 54,3 Second.

Keywords : *Hadoop, Hadoop Multi-Node, Fair Scheduling, COSHH*