Abstract

Hadoop is a java-based programming framework that supports the storing and processing of large data sets in a distributed computing environment and it is very much appropriate for high volume of data. As those reason, Hadoop frequently used for distributed processing in big data. Hadoop architecture consists of two layers which the first is layers of MapReduce and the second is layers of Hadoop Distributed File System (HDFS). MapReduce is a programming model that join to a large processing while Hadoop Distributed File System is a filesystem that the storage will save as distributed.

In MapReduce there is several Job Scheduler's algorithms that the function for data mapping, so it will improve the job quality on teh system. Hadoop clusters there is default scheduling that use FIFO algorithm. FIFO algorithm has chracteristics for every first job come will execute at the same time until finish and then continue to the next job. There are many types of scheduling in Hadoop, one of them is Self Adaptive Reduce Scheduling algorithm or as knwon as SARS algorithm. SARS algorithm performing with delay the reduce time that is use not same with jobs firts come. Delay the reduce time will decrease the response time with any jobs that will use in this scenario. SARS algorithm performance looks more effective than FIFO algorithm on the type job of wordcount with amount of jobs 50 jobs with 3,90% of fail, 157,05 minutes of response time, that resulted for job throughput 6,04 job/min better 1,93 job/min than FIFO algorithm.

Key words: *hadoop, FIFO, SARS, HDFS, jobs*