

## Abstrak

Hadoop merupakan sebuah *framework* pemrograman berbasis Java yang mendukung penempatan dan pemrosesan suatu data set yang besar dalam pendistribusian berbasis komputasi dan ini sangat sesuai dengan data dengan volume yang tinggi. Sehingga, Hadoop seringkali digunakan dalam hal pemrosesan bersifat terdistribusi terhadap big data. Hadoop memiliki dua *layer* yang pertama *layer MapReduce* dan yang kedua *layer Hadoop Distributed File System (HDFS)*. *MapReduce* merupakan suatu model pemrograman yang menggabungkan untuk suatu pemrosesan suatu data yang besar sedangkan Hadoop *Distributed File System* merupakan suatu *filesystem* yang penyimpanannya akan disimpan secara terdistribusi.

Dalam *MapReduce* ini terdapat beberapa algoritma *Job Scheduler* yang berfungsi untuk pemetaan suatu data, sehingga nantinya akan meningkatkan kualitas *job* pada sistem. Sistem Hadoop terdapat *default scheduling* dengan memakai algoritma *FIFO*. Algoritma *FIFO* memiliki karakteristik untuk setiap *job* yang akan masuk pertama akan dieksekusi langsung sampai selesai lalu dilanjutkan dengan *job* selanjutnya yang akan masuk. Algoritma *scheduling* dalam Hadoop terdapat berbagai macam, salah satunya algoritma *Self Adaptive Reduce Scheduling* atau disingkat dengan *SARS*. Algoritma *scheduling SARS* dengan karakteristiknya akan menerapkan memperlambat proses *reduce* sehingga tidak mengikuti dari awal proses *jobs* akan masuk. Proses menunda waktu *reduce* ini akan mengurangi *response time* dengan berbagai *jobs* yang akan diuji. Performansi algoritma *SARS* terlihat lebih efektif daripada algoritma *FIFO* pada jenis *job wordcount* berjumlah 50 *jobs* dengan nilai *fail* sebesar 3,90%, lama *response time* 157,05 menit mengakibatkan nilai *job throughput* sebesar 6,04 *job/menit* lebih baik 1,93 *job/menit* dibandingkan dengan *FIFO*.

**Kata Kunci:** *hadoop, FIFO, SARS, HDFS, jobs*