

## Abstrak

Hadoop merupakan *framework software* berbasis *java* dan *open-source* yang berfungsi untuk mengolah data yang besar secara terdistribusi dan berjalan diatas *cluster* yang terdiri atas beberapa komputer yang saling terhubung. Hadoop mempunyai kelebihan dari segi ekonomis karena tidak berbayar dan dapat diimplementasikan pada perangkat keras dengan spesifikasi yang tidak terlalu tinggi. Arsitektur Hadoop terdiri dari dua *layer* yaitu *layer MapReduce* dan *layer Hadoop Distributed File System (HDFS)*. *MapReduce* merupakan *framework* dari aplikasi yang terdistribusi sedangkan *Hadoop Distributed File System (HDFS)* merupakan data yang terdistribusi.

*Delay Scheduling* merupakan *job scheduler* yang sedang berkembang pada sistem Hadoop *multi-node* dan memiliki ciri dalam penanganan antrian untuk *job scheduling*. *Delay Scheduling* menerapkan metode menunda jalannya *jobs* selanjutnya untuk memperbaiki data lokalitas sebelumnya sehingga menurunkan nilai *fail* pada *job*. Selain itu, melakukan pengalokasian data yang hampir optimal sehingga berpengaruh pada *Job Fail Rate*, *Job Throughput*, dan *Response Time*. Algoritma *Delay Scheduling* memiliki performansi yang efektif dengan penurunan 0.3% *Job Fail Rate*, peningkatan 8.853% *Job Throughput*, dan lebih cepat 142 menit 45 detik *Response Time* dengan jenis karakteristik *job wordcount* pada jumlah 50 *jobs*.

**Kata Kunci:** *hadoop, hadoop multi-node, delay scheduling, FIFO.*