

Abstrak

Scheduling Hadoop merupakan cara untuk mengatur setiap *job* yang akan berjalan pada sistem Hadoop agar dapat mengelola semua *job* yang ada untuk mendapatkan giliran untuk di eksekusi pada setiap *resource* yang tersedia. Default *Scheduling* pada Hadoop yaitu FIFO yang mempunyai karakteristik untuk setiap *job* yang masuk pertama akan di eksekusi langsung dan berhak memonopoli satu *resource* secara utuh. Namun FIFO memiliki kerugian bagi proses *short job* ketika yang dieksekusi adalah proses *long job*.

Delay improve Fair share merupakan *Job scheduler* yang menggunakan metode dengan membagi *job* untuk satu cluster ke dalam beberapa *pool* dan setiap *pool* diberlakukan metode menunda jalannya *jobs* selanjutnya untuk memperbaiki data lokalitas sebelumnya. Pembagian *resource* data dan pengalokasian data yang hampir optimal akan mempengaruhi *Job Fail Rate*, *Job Throughput*, *Average Completion Time*. *Delay improve Fair share* memiliki performansi efektif daripada *Fair share* dan *Delay Scheduling* pada jenis *job* randomtextwriter terhadap data .txt dengan penurunan 0,3% *job fail rate* dengan nilai *throughput* 2,73 *job/m* dan 273,59 menit lebih cepat dari *Delay Scheduling* dan 128,15 menit lebih cepat dari *Fair share*

Kata Kunci: hadoop, hadoop multi-node, *Fair share improve Delay Scheduling improve capacity*, *Delay Scheduling improve capacity*.