

Abstrak

Proses index dokumen yang besar menghabiskan waktu yang cukup lama dalam melakukan prosesnya hingga membentuk suatu index. Diperlukan suatu indexing yang baik agar dapat melakukan indexing dengan cepat dan menghasilkan index yang tepat.

Distributed indexing merupakan suatu metode yang dapat melakukan proses index dengan cepat dikarenakan proses terjadi paralel dimana proses index dapat dilakukan di beberapa lokasi. Metode ini memiliki dua bagian utama yakni koordinator indexer dan parser. Terdapat koordinator indexer yang berguna untuk mengatur proses indexing. Koordinator Indexer mengatur dokumen yang akan di diindex, koneksi antara koordinator indexer dan parser serta mengatur inverter.

Dari implementasi yang telah dilakukan terdapat dua faktor mempengaruhi proses index yakni jumlah parser dan batas ukur pembagian dokumen yang akan diproses. Dan didapat bahwa proses lebih cepat ketika menggunakan batas ukur dan jumlah parser diperbanyak. Hasil index yang telah di proses selanjutnya akan disimpan dalam satu database.

Kata Kunci: index, distributed indexing, koordinator indexer.