

INTISARI

ANALISIS KINERJA *MAPREDUCE* UNTUK PEMROSESAN DOKUMEN PADA SISTEM *MULTICORE*

Oleh

SULVIA TRI HERAWATI

13/355447/PPA/04352

MapReduce adalah model pemrograman yang digunakan untuk memproses data dalam ukuran besar dengan menggunakan algoritma paralel dalam *cluster*. Model pemrograman ini menggunakan dua buah fungsi yaitu pemetaan atau *map* dan reduksi atau *reduce*. *MapReduce* merupakan bagian dari hadoop yang dapat digunakan di lingkungan *singlecore* dan *multicore*. Pada lingkungan *singlecore*, *MapReduce* memanfaatkan jumlah *core* yang ada pada komputer tersebut. Dengan menggunakan n *core*, kinerja pemrosesan data pada multiprosesor diharapkan meningkat n kali dibandingkan dengan kinerja pada uniprosesor.

Penelitian ini meneliti kinerja *singlecore* dan *multicore* menggunakan studi kasus *term frequency-inverse document frequency* (TF-IDF). Parameter yang digunakan pada saat pengujian berupa jumlah *file text* dan jumlah *core*. Hasil pengujian menunjukkan bahwa perbedaan kinerja *singlecore* dibandingkan *multicore* dalam pengujian studi kasus TFIDF tidak dapat terlihat dengan jelas.

Kata kunci : *MapReduce*, *Singlecore*, *Multicore*

ABSTRACT

MAPREDUCE PERFORMANCE ANALYSIS FOR DOCUMENTS PROCESSING ON MULTICORE SYSTEM

By:

SULVIA TRI HERAWATI

13/355447/PPA/04352

MapReduce is a programming model that used to process data in large sizes using parallel algorithms in *clusters*. This programming model uses two functions namely *map* and *reduce*. *MapReduce* is part of the *hadoop framework* that can be used in single node and multinode environments. In single node environment, *MapReduce* utilizes a number of *cores* available in a node. by using *n cores*, data processing performance on multi *processors* is expected to increase *n* times compared to performance on uni *processors*.

This research examined the performance of singlecore and multicore *systems* using *Term Frequency-Inverse document Frequency* (TF-IDF) case study. The parameters used during the experiments are the number of *text files* and number of *cores*. The experiment results show that the differences between singlecore performance versus multicore performance in processing TF-IDF case study can not be clearly seen.

Key word : MapReduce, Singlecore, Multicore