

ABSTRAKSI

Condro Wibawa.92214016

ANALISA PERBANDINGAN PERFORMA MAP REDUCE FRAMEWORK MENGGUNAKAN HADOOP MAPREDUCE DAN APACHE SPARK.

Tesis. Sistem Informasi Bisnis. Program Studi Magister Manajemen Sistem Informasi. Universitas Gunadarma. 2016.

Kata Kunci : Big Data, Map Reduce, Hadoop MapReduce, Apache Spark
(xiv+ 59 + lampiran)

Jumlah data yang semakin besar membuat data tidak bisa diolah dengan cara tradisional seperti yang ada saat ini. Kalaupun bisa, maka akan dibutuhkan biaya yang lebih mahal dan waktu pemrosesan yang lebih lama. Hal ini tidak sejalan dengan perkembangan dunia teknologi informasi saat ini yang menuntut proses yang cepat dengan biaya yang murah.

Salah satu cara yang dikembangkan oleh para praktisi IT adalah menerapkan model pemrograman Map Reduce. Selain menawarkan solusi pengolahan data yang sederhana dan cepat, kelebihan lain dari model pemrograman ini adalah bisa dijalankan dalam komputer cluster secara paralel.

Saat ini ada beberapa framework Map Reduce yang bisa digunakan. Dua diantaranya adalah Hadoop MapReduce dan Apache Spark. Kedua framework ini memiliki konsep dan cara kerja yang sama. Perbedaannya adalah pada penggunaan resource penyimpanan data sementara. Hadoop MapReduce menggunakan HDFS (Disk) sebagai media penyimpanan data, Sedangkan Apache Spark menggunakan RDDs (in-memory).

Penggunaan RDDs oleh Apache Spark membuat kinerja framework ini jauh lebih cepat dibandingkan dengan Hadoop MapReduce. Terbukti dalam penelitian ini rata-rata kinerja Apache Spark 4,99 kali lebih cepat dibandingkan dengan Hadoop MapReduce.

Daftar Pustaka (2004 – 2016)