

Abstrak

Pengolahan data dari sebuah sistem *Polling* menjadi hal yang sangat penting karena data hasil olah tersebut dapat digunakan oleh publik untuk dijadikan acuan di dalam menghadapi sebuah persoalan yang berkembang di masyarakat. Pertumbuhan di dalam penggunaan media sosial semakin meningkat dari tahun ke tahun, di mana Indonesia sendiri merupakan pengguna media sosial *Twitter* terbanyak ke-5 di dunia. Di dalam proses pengolahan data, jika data yang akan diolah berukuran cukup besar, akan memperlambat komputasi yang dilakukan. Hal tersebut mendorong penulis untuk membuat sebuah sistem yang dapat memproses data *polling* yang dilakukan melalui media sosial dengan waktu yang lebih efisien. *Hadoop* merupakan salah satu sistem yang optimal untuk digunakan di dalam pengolahan data *polling* pada Tugas Akhir ini. Pada *Hadoop* terdapat 2 modul utama yaitu *Hadoop Distributed File System (HDFS)* yang merupakan sistem penyimpanan terdistribusi, dan *MapReduce* yang merupakan algoritma/komputasi pada *Hadoop*. Pada pengolahan data *polling* ini menggunakan program *wordcount* dengan *MapReduce* pada *Hadoop* dan dengan program *wordcount* tanpa *MapReduce*. Dilakukan pengujian terhadap 2 metode tersebut, dengan diujikan menggunakan beberapa data dengan ukuran dari yang kecil sampai ke yang berukuran besar. Dan dihasilkan bahwa, *MapReduce* lebih unggul dalam segi kecepatan proses data dibandingkan dengan metode proses data tanpa *MapReduce*. Dengan rata-rata dari data yang diujikan, menggunakan *MapReduce* pada *Hadoop* dapat memproses data 1,3 kali lebih cepat dibandingkan tanpa *MapReduce* pada *Hadoop*.

Kata Kunci: *polling, Hadoop, MapReduce, wordcount, kecepatan proses.*