

Abstrak

Hadoop adalah *software framework opensource* yang memungkinkan untuk pemrosesan secara terdistribusi dari kumpulan set data yang besar pada *Cluster* komputer dengan menggunakan model pemrograman tertentu. Hadoop terdiri dari 2 buah komponen utama dalam menyelesaikan permasalahan *Big Data*, yaitu MapReduce dan HDFS (*Hadoop Distributed File System*).

Dari kedua komponen tersebut maka HDFS yang memegang peranan penting dalam mengatur *File System*. Namun saat *MasterNode* yang di dalamnya terdapat *Namenode* dan *Jobtracker* mengalami gangguan maka *File System* tidak dapat melayani pengguna. Maka dari itu, diperlukan cara khusus untuk dapat mengatasinya dengan membuat *cloning* dari *MasterNode* di server yang lain sehingga saat *MasterNode* mengalami masalah dapat digantikan oleh *cloning*-nya yang disebut dengan *StandbyNode*.

Pada pengerjaan Tugas Akhir ini menerapkan metode Zookeeper dan QJM pada *cluster* Hadoop untuk dapat meningkatkan *High Availability* pada Hadoop dengan *downtime failover* berkisar antara 4 detik sampai 63 detik. Jika dibandingkan dengan metode DRBD dan Heartbeat maka metode Zookeeper dan QJM yang digunakan sekarang lebih baik untuk digunakan dalam jangka panjang. Dan dengan menggunakan Zookeeper dan QJM, *downtime failover* relatif stabil tanpa adanya peningkatan *downtime* terhadap jumlah blok data.

Kata kunci : Hadoop, High Availability, NameNode, Failover, Downtime