

## ABSTRACT

Yandi Eko Testiono, 92215064.

### **“IMPLEMENTATION OF HADOOP FRAMEWORK ON MANAGEMENT OF BIG DATA ARSIP NASIONAL RI.”**

Thesis. Program Pascasarjana, Universitas Gunadarma, 2017.

Keywords: Big Data, Hadoop Framework, Volume, Variety, and Velocity.  
(xv + 114 + Attachment)

Big Data Technology is a high-volume, complex information management asset that assists companies in performing cost-effective data management, and as decision-making. Big Data ensures the processing of data solutions with several new and existing variants to deliver tangible benefits in the business world. This study aims to solve the problem of the Arsip Nasional RI Institution in handling large amount of data (Volume), Rapocity growth, and Variety formats during filing of data archives. The method in this study is approach through Hadoop framework, and based on related journal literature. The results of this study explain the implementation Hadoop *single node* using two methods: manual implementation, and implementation using Big Data Hadoop service. The first method is manual implementation using the "Cygwin, VMWare, Eclipse, Web Browser" method, and the "Command Prompt, Web Browser" method. The second method is to use the Big Data Hadoop service using Cloudera. This research proves Big Data implementation has addressed the problems occurring in Arsip Nasional RI Institute, such as *storage problems*, and *Disaster Recovery Plan*. This is because Hadoop has a management and operational structure that is provided in the form of an open and integrated service platform that provides affordable storage and computing options when viewed in terms of cost and efficiency. In addition Hadoop also provides *Disaster Recovery Plan* feature in the form of Cloudera service, which allows for data replication across data centers. This replication includes data stored in HDFS, data in the Hive table, Impala metadata listed in Metastore Hive. When important data is stored in HDFS, Cloudera Manager helps ensure data is available at all times, even when there is a stop in the data center.

Bibliography (2009 - 2016)