

ABSTRACT

Big data to popular for audience computer science around the world .So much data with large amounts and bervariatif be so difficult in an analysis data .In processing big data mining will different from the data mining. Big data mining with the amount of data very large to be done as soon as possible with broken down to some node (parallel computing) to get the information which is useful. For it will be analyzed to determine the processing of big data mining. The analysis is done through simulation with multi-node configuration which consists of multiple computers that have been configured. Big data mining in this final project will use clustering method. In this clustering method the algorithm used is K-means. The result of big data clustering means obtaining an optimal cluster of $K = 5$ and the timing depends on cluster mode, number of worker and hardware specs used, if low-specification hardware is used when the automatic cluster mode capability for processing follows such low-specification hardware.

Keywords: *Big Data, Data Mining, Clustering, K-Means, Cluster Mode*