

Abstract

A large document indexing process consume a lot of time when doing its process until an index is constructed. A good indexing is required to perform a fast indexing while creating the right index.

Distributed indexing is a method that can be perform a fast indexing because the indexing process is done parallel whereas the indexing process is done on a several location. This method has two main parts which are the koordiantor indexer and the parser. There is an indexer coordinator that controls the indexing process. The coordinator controls the documents which are going to be indexed, the connection between indexer coordinator and parser, and controls the inverter.

From the implementation that had been done, there are two factors that affects the indexing process which are the number of parser and the threshold of the document size that are going to be processed. The process is also faster if using threshold and a larger number of the parser. The index that had been processed will be stored in a single database.

Key Word: *index, indexer coordinator, distributed Indexing*