

Abstract

There are two flaws in massive graph which is implemented using only one computer is the problem of search time and memory limitations. Each computer has limited memory so massive implementation of a data graph can degrade computer performance. Both deficiencies previously described can be overcome by the distribution of the data.

The distribution of data is done by partitioning the existing graph database. There are various techniques for graph database partition. In this thesis, the partition graph database is implemented using the engineering division or partition graph by using minimum communication method partitioning (MCP). MCP is a method of partitioning algorithms based Distributed Breadth-First Search (dBFS). Distribution of data is done on a shared-nothing parallel system.

There are three functionalities on systems built with the distribution of the data is to look for the presence of neighborhood vertex layout (pattern), look for the neighbors on each node (cross-partition), and look for the search time based id vertex on each node. The aim of this thesis was to determine how to implement data distribution using MCP partition technique and to determine the performance of query response time in a system that implements data distribution and systems that do not implement data distribution

Keywo: *Graph Database, distribution of the data, partition graph database, massive graph, DBFS, MCP*