Abstract

The performance of a parallel program depends largely on its data partitions. So a good data partitioning scheme is the need of the time. However, it is very difficult to arrive at a good solution as the number of possible data partitions for a given real-life program is exponential in the size of the program. In this paper, we present a heuristic technique for automatic data partitioning for HPF. Our approach is based on Genetic Algorithms and is very simple, yet very efficient to quickly find appropriate data partitions even for large programs with large number of alternatives for data distribution. It makes use of both static as well as dynamic data distribution with the main aim of reducing the overall execution time of the entire program. Experimental results based on tests performed on a 4-node and a 8-node Linux cluster are also shown.