Balance Algorithm - a New Approach to Solving the Mapping Problem on Heterogeneous Systems

Nguyen Thanh Son, Tran Nguyen Hoang Huy, and Nguyen Anh Kiet

Department of Information Technology, HCM University of Technology, Vietnam
sonsys@dit.hcmut.edu.vn
huy@dit.hcmut.edu.vn

Summary. A fundamental issue affecting the performance of a parallel program is the assignment of tasks to processors in order to achieve the minimum completion time. Most of state-of-the-art approaches consider homogeneous MIMD multiprocessor systems, in which all communication channels have the same bandwidth and all processors are equally powerful. These algorithms do not run efficiently on heterogeneous systems. In this paper, we present a new approach for the mapping problem on arbitrary systems. The main idea is based on the "global load balancing, local cut-size optimization" principle. This approach has achieved encouraged results that are verified by experiments for various random graphs and processor numbers.

Key words: partitioning, mapping, load balance, heterogeneous system

1 Introduction

One of the most important concerns in parallel computing is the proper distribution of work-load across processors, namely the mapping problem. The problem amounts to the balance of the computational weight on the processors and reduces the communication cost by keeping intensively intercommunicating processes on nearby processors.

A mapping is static if it is computed prior to the execution of the program and is never modified at runtime. Static mapping is NP-complete in general case. Therefore, numerous methods have been developed to solve heuristically the mapping problem for bus, ring, mesh and hypercube. If the communication network of the parallel system is a complete graph, then the mapping problem turns into the graph partitioning problem. There are many popular graph partitioning approaches such as Kernighan-Lin [11], Fiduccia-Mattheyes [2], Spectral Bisection [4], Genetic [1], Simulated Annealing [13], [6], etc. Most of them have investigated homogeneous systems, in which all communication channels have the same bandwidth and the processors are equally powerful.
Thus, it is not efficient to use these algorithms on heterogeneous systems without necessary improvements. Popular partitioning packages are METIS (by G. Karypis and V. Kumar at University of Minnesota) [10], and Chaco (by B. Hendrickson and R. Leland at Sandia National Lab.) [5]. Both of them use multilevel Kernighan-Lin method. One prominent feature of METIS is that it can handle the case of nonuniform partition sizes. However, it only consider the processor powers. The communication links have not yet been taken into account. Hence, traditional graph partitioning algorithms are not adequate to ensure the efficient mapping solutions for heterogeneous systems.

Based on the ideas of Kernighan-Lin and Fiduccia-Mattheyses algorithm, we combine their advantages to present an effective algorithm for heterogeneous systems. This algorithm, namely Balance Mapping, can be used for mapping a parallel program onto a parallel system. Furthermore, we proposed a flexible objective function that overcomes the disadvantages of many mapping packages to handle the case of heterogeneous systems.

The rest of the paper is organized as follows. In Section 2, we summarize recent mapping solutions. Section 3 focuses on our proposed objective function for heterogeneous systems. The main idea of new algorithm, namely Balance, is presented in Section 4. Experimental results, including the execution time, the edge-cut ratio and load imbalance ratio on several task graphs and system graphs are showed in Section 5. Then follows the conclusion.

2 Preliminaries

2.1 Overview and related works

A parallel program consists a number of tasks that run concurrently. These tasks might have to communicate with each other by sending or receiving messages. We model the parallel program as a weighted graph $S(V, E)$, namely task graph. Vertices $v_S$ and edges $e_S$ of $S$ are assigned integer weight $w(v_S)$ and $w(e_S)$, which estimate the computation weight of the corresponding process and the amount of communication on the inter-process channels, respectively. In this paper, we consider the mapping problem for heterogeneous multi-computers. Message passing is used for communication between computers. The parallel system is modelled by a weighted graph $T$ called system graph. Vertices $v_T$ and edges $e_T$ of $T$ are assigned integer weight $w(v_T)$ and $w(e_T)$, which estimate the computational power of the corresponding processor and the cost of the inter-processor links, respectively.

Mapping a task graph onto a system graph is a function $M : S \rightarrow R$ such that $M(i)$ gives the processor onto which the task $i$ is mapped. The main objective is to find an $M$ that minimizes the overall execution time of the parallel program. Numerous methods have been developed to solve the mapping problem. The blind-search methods of artificial intelligence-breadth-first or depth-first-are exhaustive methods for finding an optimal mapping.