A Performance Analysis of Single versus Multiple Processors

E.Arthurs

Bell Laboratories
Murray Hill, New Jersey 07974

B.W.Stuck

Bell Laboratories
Murray Hill, New Jersey 07974

ABSTRACT

We consider a simple model of a multiprocessor digital system, with three main ingredients: a load of different transaction types with given arrival statistics, an architecture consisting of a network of loosely coupled nodes with tightly coupled symmetric multiple processors at each node, and a load partitioning or assignment of each step of execution of each transaction type to a given node. In order to quantify this problem, we adopt a Jackson network for ease of analysis. Our goal is to consider tradeoffs between the workload, the architecture, and workload partitioning to maximize performance. We measure performance by two indices, the mean throughput rate of completing each transaction type, and the mean response time to completely execute each transaction type.

We present analytic formulae for the mean response time for a job assuming simple Poisson arrival statistics.

To illustrate the many tradeoffs possible, we consider one example in detail. We show that if we fix the total mean instruction rate, the mean response time for the multiprocessor network is always greater than the mean response time for a job submitted to one equivalent fast processor. In the special case of purely random arrivals and equal utilization of all network nodes, the mean response time will be at least N times that of a single equivalent fast processor. Moreover, the optimum assignment of excess processing capacity to each node to minimize mean response time does not obey a simple linear proportionment based on utilization, but rather a square root assignment of excess capacity.

This model accounts for randomness in arrival and service patterns. Significant features left out of this analysis are the system overhead in coordinating all the multiprocessor interactions, memory contention, and so forth, which presumably will only further degrade performance. In many applications, other issues, such as flexibility, reliability, cost, and so forth, may play a dominant role in a design decision; here we concentrate solely on the system response time as the performance measure, with a fixed total instruction rate.