Performance Prediction and Procurement in Practice: Evaluating Commodity Cluster Components for Wavefront-Heavy Workloads

S.D. Hammond

High Performance Systems Group
Department of Computer Science
University of Warwick
Coventry, CV4 7AL, UK

Abstract. The cost of state-of-the-art supercomputing resources make each individual purchase an expensive and often lengthy process. In selecting a machine, the procurement manager must balance complex user requirements against organisation objectives, machine size against cost and individual component selection against intended performance. In traditional HPC purchases benchmarking of small representative applications is used to highlight suitable architectures. This however provides insights limited to the selected benchmark, specific problem inputs and restricted processor counts. Recent developments in modelling the performance of large industrial applications can help to alleviate some of these issues by providing a largely automated set of methods for exploring code behaviour on machines consisting of thousands of processors.

In this talk we present the application of the recently developed WARwick Performance Prediction (WARPP) toolkit in evaluating a set of commodity components which may be combined in a future mid-range cluster purchase. Our focus is limited to processor core counts of less than 4096 - a size typical of University and large organisation purchases. Our models, which demonstrate predictive accuracies of greater than 90%, are used to provide insight into the code performance that results from choices in networking interconnects, individual processor performance and higher core densities.

The contributions of the work presented are to explore a set of complex machine parameters which directly impact the performance of wavefront codes at various processor counts. In analysing our results we are able to demonstrate the compute bound nature of these algorithms at lower processor counts and the heavy reliance on low-latency interconnects when higher core counts are employed. These results may help to direct future computing machinery purchases or upgrade component selection in the context of wavefront-heavy workloads.