Workshop 2+8
Performance Evaluation and Prediction

Allen D. Malony and Rajeev Alur

Co-chairmen

Measurement and Benchmarking

The core of any performance evaluation strategy lies in the methodologies used for measuring performance characteristics and comparing performance across systems. The first paper in this session describes an efficient performance measurement system created for workstation clusers that can report runtime information to application level tools. A hardware monitor approach is argued for in the second paper for measuring the memory access performance of a parallel NUMA system. The third paper demonstrates analysis methods for testing self-similarity of workload characterizations of workstations, servers, and the world-wide web. It is hard enough deciding performance metrics for parallel application benchmarking, but making the methodology portable is a real challenge. This is the subject of the fourth paper.

Modeling and Simulation

Parallel systems pose unique challenges for constructing analytical models and simulation studies because of the inter-relationship of performance factors. The first paper deals with the problem of building probabilistic models for CPU's with multilevel memory hierarchies when irregular access patterns are applied. The second paper studies the validity of the h-relation hypothesis of the BSP model and proposes an extension to the model based on an analysis of message passing programs. Simulating large-scale parallel programs for predictive purposes requires close attention to the choice of program factors and the effect of these factors on architectural analysis. This is shown effectively in the third paper for predicting data paralel performance.

Communications Modeling and Evaluation

Scalable parallel systems rely on high-performance communication subsystems and it is important to relate their characteristics to communication operations in software. The first paper assess LogP parameters for the IBM SP machine as applied to the MPI library and distinguished between sender and receiver operations. In contrast, pre-evaluating communications performance at the level of a data parallel language, independent of machine architecture, relies instead

on the aspects of data distribution and communication volume and pattern. The second paper pre-evaluates communication for HPF. Targeting both a communcations model and its application to program development, the third paper applies genetic programming to the problem of finding efficient and accurate runtime functions for the performance of parallel algorithms on parallel machines.

Real-Time Systems

The real-time systems session is dedicated to the design, implementation, and verification techniques for computerized systems that obey real-time constraints. The subject comprises programming languages for real-time systems, associated compiling or synthesis technology, real-time operating systems and schedulers, and verification of real-time constraints. The first paper addresses the issue of language support for design of reliable real-time systems, proposing a methodology based on constraint programming. The second paper presents theoretical results concerning scheduling of periodic processes. In particular, the authors present sufficient conditions, weaker than previously known, for guaranteed schedulability so as to meet given age constraints. The third paper describes how the authors formulated and solved a control problem in the design of an industrial grinding plant to optimize economic criteria using parallel processin.