Guest Editorial: Special Section on Algorithm Design and Scheduling Techniques (Realistic Platform Models) for Heterogeneous Clusters

Henri Casanova, Member, IEEE, Yves Robert, Fellow, IEEE, and H.J. Siegel, Fellow, IEEE

The last decade has seen a dramatic increase in the deployment of heterogeneous distributed computing platforms, in particular, those consisting of heterogeneous clusters, and multiple heterogeneous collections of clusters aggregated over wide-area networks into grids. The software infrastructures and mechanisms to deploy such platforms have been well studied and implementations are already used in production, so that heterogeneous platforms represent a significant, and growing, fraction of the computational power delivered by parallel platforms today. In spite of these successes, many research challenges remain, including those pertaining to distributed algorithms and scheduling algorithms, which are critical for ensuring that these platforms are used effectively. In this context, the goal of this special section on “Algorithm Design and Scheduling Techniques (Realistic Platform Models) for Heterogeneous Clusters” is to gather papers that further our understanding of the impact of platform heterogeneity on the design and evaluation of new such algorithms.

In the paper entitled “Allocating Non-Real-Time and Soft Real-Time Jobs in Multiclusters,” Ligang He, Stephen A. Jarvis, Daniel P. Spooner, Hong Jiang, Donna N. Dillenberger, and Graham R. Nudd introduce two workload allocation strategies for large-scale heterogeneous platforms. The first strategy achieves an optimized mean response time for jobs having no real-time requirements. The second strategy achieves an optimized mean miss rate for jobs having soft real-time requirements (i.e., a fraction of jobs are permitted to miss the real-time constraints). Both strategies take into account average system behaviors (such as the mean arrival rate of jobs) to calculate the workload proportions for individual clusters, and update on-the-fly the workload allocation when the change in the mean arrival rate reaches a certain threshold. The allocation schemes are combined with two job dispatching strategies (weighted random and weighted round-robin) to generate new job scheduling algorithms for multicluster environments.

In their paper “On the Distribution of Sequential Jobs in Random Brokering for Heterogeneous Computational Grids,” Vandy Berten, Joel Goossens, and Emmanuel Jeannot study resource brokering for scheduling sequential jobs onto a grid platform that consists of heterogeneous sets of homogeneous processors, such as a set of clusters. Resources in each cluster are managed by a local scheduler that maintains a job queue. The paper studies a centralized “metascheduler” that uses a randomized strategy to share available resources among competing jobs. This research considers two cases depending on whether the platform is heavily loaded or lightly loaded. For each case, it obtains both analytical and experimental characterizations of the queue lengths at each local scheduler, CPU utilization, and average job slowdowns. Furthermore, the paper presents a discussion of the system’s behavior when it transitions between a heavily loaded state and a lightly loaded one. All presented theoretical results are corroborated by simulations and provide a thorough description of randomized resource brokering.

The research in “Multiple Job Scheduling in a Connection-Limited Data Parallel System” presents a new method for scheduling jobs in a distributed system where the critical resource is the bandwidth to access the stored data. The authors, Alessandro Amoroso and Keith Marzullo, describe an approach that supports the master-worker scheme and can be applied to data parallel computation. They consider a typical wide-area data grid that is comprised of a set of sites, where each site has one or more local area networks. The platform model used is based on the Nile data grid. This paper uses a set of synthetic jobs to compare three schedulers: Greedy, Maxflow, and Hybrid. They tested their new approach under various circumstances and measured its performance by means of several metrics. The new Hybrid scheduler is never worse than either of the other two schedulers, and in 20 percent of the simulated runs, it produced runs that were at least 20 percent better.

The paper entitled “Capacity-Aware Multicast Algorithms on Heterogeneous Overlay Networks,” coauthored by Zhan Zhang, Shigang Chen, Yibei Ling, and Randy Chow, addresses the problem of multicast for group
communication among a distributed, dynamic set of heterogeneous nodes. Two capacity-aware overlay multicast services that focus on host heterogeneity, any-source multicast, dynamic membership, and scalability are proposed. Capacity is modeled as the maximum number of direct children to which a node is willing to forward multicast messages. The target applications considered are multisource environments, such as distributed games, teleconferencing, and virtual classrooms. They extend Chord and Koorde to be capacity-aware, and embed implicit degree-varying multicast trees on top of the overlay network and develop multicast routines that automatically follow the trees to disseminate multicast messages. They analyze the expected performance of the proposed multicasting schemes and perform simulations. The simulations show that the two methods achieve their best performances under different conditions, depending on membership change frequencies and node capacities.

We are very grateful to all who have helped in bringing about this special section. Twenty-one papers were submitted, of which only four were accepted. We thank the authors of all submitted papers for their interest in this special section, as well as the reviewers for their insightful comments and recommendations. We wish to acknowledge the excellent job of Ms. Suzanne Werner and Ms. Jennifer Carruth for helping manage the entire submission, review, and publication process. Finally, we thank Dr. Pen-Chung Yew, the former Editor-in-Chief of IEEE Transactions on Parallel and Distributed Systems, who originally envisioned a special issue on algorithms and scheduling, and who proposed the idea at the TPDS editorial meeting held during IPDPS 2004 in Santa Fe, New Mexico.

Henri Casanova
Yves Robert
H.J. Siegel
Guest Editors

Henri Casanova received the BS degree from the Ecole Nationale Superieure d’Electronique, d’Electrotechnique, d’Informatique et d’Hydraulique de Toulouse, France in 1993, the MS degree from the Universite Paul Sabatier, Toulouse, France in 1994, and the PhD degree from the University of Tennessee, Knoxville in 1998. He is an assistant professor of computer and information sciences at the University of Hawaii at Manoa. His research interests are in the area of parallel and distributed computing, with a focus on modeling, simulation, and scheduling. He is a member of the IEEE.

Yves Robert received the PhD degree from the Institut National Polytechnique de Grenoble in 1986. He is currently a full professor in the Computer Science Laboratory LIP at ENS Lyon. He is the author of four books, 95 papers published in international journals, and 115 papers published in international conferences. His main research interests are scheduling techniques and parallel algorithms for clusters and grids. He is a fellow of the IEEE and the IEEE Computer Society, and serves as an associate editor of IEEE Transactions on Parallel and Distributed Systems.

H.J. Siegel received two BS degrees from the Massachusetts Institute of Technology, and the MA, MSE, and PhD degrees from Princeton University. He was appointed the George T. Abell Endowed Chair Distinguished Professor of Electrical and Computer Engineering at Colorado State University (CSU) in August 2001, where he is also a professor of computer science. In December 2002, he became the first director of the university-wide CSU Information Science and Technology Center (ISTeC). From 1976 to 2001, he was a professor at Purdue University. Professor Siegel has coauthored more than 300 published papers on parallel and distributed computing and communication. He is a fellow of the IEEE and a fellow of the ACM. He was a Coeditor-in-Chief of the Journal of Parallel and Distributed Computing, and was on the Editorial Boards of both the IEEE Transactions on Parallel and Distributed Systems and the IEEE Transactions on Computers. He was program chair/cochair of three major international conferences, general chair/cochair of six international conferences, and chair/cochair of five workshops. Professor Siegel is chair of the Steering Committee for the annual IEEE Heterogeneous Computing Workshop. He has been an international keynote speaker and tutorial lecturer, and has consulted for industry and government. For more information, please see www.engr.colostate.edu/~hj.

For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.