Editorial Paper

CURRENT TRENDS IN MODELLING AND SIMULATION OF PARALLEL AND DISTRIBUTED SYSTEMS

HELEN D. KARATZA

Department of Informatics
Aristotle University of Thessaloniki
54124 Thessaloniki, GREECE
Email: karatza@csd.auth.gr

Parallel and distributed Systems have considerable computational power which can be used to solve problems with large computational requirements.

Many different types of systems are available, some consisting of large numbers of processors. Some of the possible types are traditional multiprocessor vector systems, shared memory multi-threaded systems, distributed shared memory MIMD systems, distributed memory MIMD systems, and clusters of PCs or workstation networks.

The availability of parallel and distributed systems as well as the diversity of hardware and software makes the arbitration and management of resources among users very difficult. It is crucial to decide where and when an application should execute, i.e., the processor an application processes (or threads) will run on and the order they will run.

Parallel and distributed computing is valuable for problem solving in many scientific domains and has gained a large amount of attention. Performance studies have explored a wealth of alternatives with respect to parallel and distributed systems aimed at improving system performance by improving the performance of its individual components. Other research efforts have focused on improving networks in order to deliver high performance. The References Section of this paper lists only a few examples from the literature.

The most straightforward way to evaluate the performance without building a full-scale implementation is through modeling and simulation. Simulation models can help determine performance bottlenecks inherent in an architecture and provide a basis for refining the system configuration.

The objective of this special issue is to present recent high quality papers that deal with Modeling and Simulation of Parallel and Distributed Systems. This issue contains eight regular papers that cover several important areas of parallel and distributed systems. The papers are representative of the significant research efforts being performed worldwide. They provide insight into current and future trends of modeling and simulation in parallel and distributed systems.

Boykin's and Znati's paper deals with packet scheduling research in high speed integrated, multimedia services networks (ISNs) that aims to efficiently emulate the Generalized Processor Sharing (GPS) service discipline. Their paper presents a new scheduling approach, QoS-Aware Fair Sharing (QFQ), designed to reduce packet scheduling overhead while guaranteeing the QoS requirements of supported applications. Because the central goal of the QFQ discipline lies in meeting the QoS requirements of backlogged flows and not guaranteeing a weighted-fair portion of the server's output to such flows, the QFQ approach provides ISN server's with a great deal of flexibility in determining how to service contending application flows. Their paper presents this new and innovative scheme which seeks to exploit an excess capacity allocation policy in order to achieve a reduction in service overhead. The paper describes how such an objective can be achieved and shows how excess capacity can be identified for this purpose despite the presence of non-empty service queues.

Gerasimov's and Simon's paper describes QoS-AODV¸ an integrated End-to-End Delay (EED) constrained route discovery and bandwidth reservation protocol. QoS-AODV is designed to operate within a TDMA network. Unlike other path finding protocols that ignore the impact of the data link layer, QoS-AODV incorporates slot-scheduling information to ensure that bandwidth reserved and EED requirements satisfied as well. To minimize EED, QoS-AODV uses an efficient heuristic algorithm for TDMA slot assignment that in addition provides zero delay jitter. QoS-AODV is an enhanced version of the Ad hoc On-demand Distance Vector routing protocol, and is therefore compatible with proposed route discovery and maintenance techniques. In order to test the effectiveness of this protocol the authors implemented a version of QoS-AODV, with several different slot-scheduling approaches, in the ns-2 simulator. Their experiments showed the QoS-AODV is capable of finding the high quality traffic routes necessary to support highly synchronized parallel and distributed applications.

Ashouei et al. working closely with subsurface sensing and imaging (SSI) researchers, have developed a number of techniques that allow them to quickly characterize an application's serial execution. Their goal is to effciently map large applications and data sets on to both Beowulf-class clusters and shared memory multiprocessor systems. They utilize MPI to support a message passing, distributed processing, model. They describe their techniques, illustrating how they can quickly develop parallel implementations of serial applications. In their paper they present two example SSI applications to demonstrate the power of their techniques. They have selected to parallelize a Monte Carlo Simulation of Acousto-Photonic Imaging (API) and the Steepest Descent, Fast Multipole Method (SDFMM).

Kouvatsos et al. present a novel and efficient analytic framework for the performance analysis and capacity-assignment optimisation of a wireless GSM cell employing the RUP (Re-Use Partitioning) policy. RUP splits hierarchically the available bandwidth into multiple layers of frequencies and allows tighter frequency re-use in order to achieve a higher network capacity. In this context, they propose a queueing network model (QNM) of a wireless GSM cell consisting of a hierarchical layer configuration which is decomposed into individual GE/GE/c/c loss systems each of which is analysed in isolation via a more general maximum entropy state probability solution, subject to appropriate GE-type flow formulae and mean value constraints. Moreover, they propose a new performance optimisation index (POI) as the weighted average non-blocking probability of traffic over all frequency layers. For illustration purposes, they utilized the POI to formulate and solve two capacity-assignment optimisation problems. They included numerical examples to validate the relative accuracy of the analytic GE-type performance metrics against simulation and assess the optimal re-use partitioning policy of the available bandwidth.

He, Tian and Chen in their paper discuss the reliability of distributed systems in which nodes may fail with certain probabilities. The distributed systems have been modeled by a probabilistic graph G. They focus on the communication capacity that is characterized by a particular reliability attribute, the residual connectedness reliability, denoted by R(G). The residual connectedness reliability is the probability of all the residual nodes being reachable to each other. They first propose a deterministic bounding approach to bound R(G). They obtain a deterministic upper bound and a lower bound. To prove that their bounds are tight, they demonstrate theoretically and numerically that the difference between the upper and the lower bounds gradually tends to zero as the number of nodes tends to infinite under the condition that the node failure probability is reasonably low. That is, for large distributed systems the upper and lower bounds give an accurate estimation of R(G). They show that this approach doesn't work well for small and middle-sized systems with a high failure probability. In the second part of the paper, they present new approach that combines a Monte Carlo simulation scheme and their deterministic bounding approach to obtain a probabilistic point estimator for R(G). They also determine the confidence interval of the estimator.

Zuberek in his paper studies multiprocessor systems with different numbers of processors but with the same utilizations of corresponding components; such systems are called performance equivalent. Performance equivalence can be used to simplify simulation-based performance analysis of complex systems by simulating much simpler systems which are equivalent with respect to performance to the original ones. In this paper it is shown that in some cases identifying performance equivalent systems is quite straightforward.

Kabalan, Smari and Hakimian in their paper discuss several adaptive load sharing algorithms for heterogeneous distributed computing systems, and they propose some modifications to existing algorithms. The modified solutions are simulated to compare their processing time, resource allocation, communication costs and sensitivity to parameter changes. They conclude from the results of the simulation runs that simple algorithms, that do not require much computational overhead or implementation complexity, can perform very well under most operating conditions

Maqousi and Ball present the development of a measurement based flow acceptance control mechanism, which is used in conjunction with Class Based Queuing (CBQ) to support guaranteed services in multiservice packet switched networks. Using simulation models they present simulation experiments with monitoring techniques that monitors the mean and variance of packet inter-arrival times and packet lengths of a continuous media packet stream. M-FAC bases admission decision on the parameters provided by source, however, once the new source has been accepted, it monitors the arriving process to determine the true load vector, which is used in subsequent admission test. They describe the monitoring process and explain how M-FAC uses this to adjust its estimate of the instantaneous load vector, and present results from experiments to test M-FAC responsiveness to dynamically changing traffic load.

ACKNOWLEDGMENT

I would like to thank all authors who submitted papers, including those whose papers were not selected for this Special Issue. A special note of thanks goes to the anonymous reviewers for their abundant time and effort. Without them, this issue would not be possible. Thanks are also due to the Editor-in-chief Dr. David Al-Dabass for his help, constructive suggestions, solid support and excellent cooperation.

REFERENCES

Dandamundi S. 1994. "Performance Implications of Task Routing and Task Scheduling Strategies for Multiprocessor Systems". In *Proc. of the IEEE-Euromicro Conf. on Massively Parallel Computing Systems* (Ischia, Italy, May) IEEE Computer Society, Los Alamitos, CA, USA. Pp348-353.

Dandamudi S. 1997. "The Effect of Scheduling Discipline on Dynamic Load Sharing in Heterogeneous Distributed Systems". In *Proc. of the 5th International Workshop on Modeling, Analysis, and Simulation of Computer and Telecommunications Systems* (Haifa, Israel, January) IEEE Computer Society, Los Alamitos, CA, USA. Pp17-24.

Feitelson D.G. and Rudolph R. 1995. "Parallel Job Scheduling: Issues and Approaches". In *Job Scheduling Strategies for Parallel Processing*, *Lecture Notes in Computer Science*, Springer-Verlang, Berlin, Germany. Vol. 949. Pp1-18.

Feitelson D.G. and Rudolph L. 1996. "Evaluation of Design Choices for Gang Scheduling Using Distributed Hierarchical Control". *Journal of Parallel and Distributed Computing*, Academic Press, New York, USA, Vol. 35. Pp18-34.

Feitelson D.G. and Jette M.A. 1997. "Improved Utilisation and Responsiveness with Gang Scheduling". In *Job Scheduling Strategies for Parallel Processing*, *Lecture Notes in Computer Science*, Springer-Verlang, Berlin, Germany. Vol. 1291. Pp238-26.

Karatza H.D. and Hilzer. R.C. 1999. "Processor Failures in a Shared - Memory Multiprocessor System with Resequencing". In *Proc. of the High Performance Computing Symp. '99* (San Diego, CA, April) SCSI, San Diego, CA, USA. Pp243-248.

Karatza H.D. 1999a. "A Simulation-Based Performance Analysis of Gang Scheduling in a Distributed System". In *Proc. of the 32nd Annual Simulation Symp.* (San Diego, CA, USA, April) IEEE Computer Society, Los Alamitos, CA, USA. Pp26-33.

Karatza, H.D. 1999b. "Gang Scheduling in a Distributed System with Processor Failures". In *Proc. of the UK Performance Engineering Workshop* (Bristol, UK, July) University of Bristol, Bristol, UK. Pp199-208.

Karatza H.D. 2000a. "Gang Scheduling and I/O Scheduling in a Multiprocessor System". In *Proc. of 2000 Symp. on Performance Evaluation of Computer and Telecommunication Systems* (Vancouver, Canada, July) SCSI, San Diego, CA, USA. Pp245-252.

Karatza H.D. 2000b. "A Comparative Analysis of Scheduling Policies in a Distributed System using Simulation", *International Journal of SIMULATION Systems, Science & Technology*, UK Simulation Society, UK, Vol. 1 (1-2). Pp12-20.

Karatza H.D. and Hilzer R.H. 2001. "Epoch Load Sharing in a Network of Workstations". In *Proc. of the 34th Annual Simulation Symposium* (Seattle, WA, April) IEEE Computer Society, Los Alamitos, CA, USA. Pp36-42.

Karatza H.D. 2001. "Epoch Scheduling in a Distributed System". In *Proc. of the Eurosim 2001 Congress* (Delft, Netherlands, June) Eurosim, Delft, Netherlands. Pp1-6.

Maheswaran A.M.S., Siegel H.J., Hensgen D. and Freund R.F. 1999. "Dynamic Matching and Scheduling of a Class of Independent Tasks onto Heterogeneous Computing Systems". In *Proc. of the Eighth Heterogeneous Computing Workshop* (San Juan, Puerto Rico, April) IEEE Computer Society, Los Alamitos, CA, USA. Pp30-44.

Rosti E., Serazzi G., Smirni E. and Squillante M. 1998. "The Impact of I/O on Program Behavior and Parallel Scheduling". *Performance Evaluation Review*. ACM, New York, USA. Vol. 26 (1). Pp56-65.

Rosti E., Smirni E., Serazzi G. and Dowdy L.W. 1995. "Analysis of Non-Work-Conserving Processor Partitioning Policies". In *Proc. of the Workshop on Job Scheduling Strategies for Parallel Processing* (Santa Barbara, CA, April) IEEE Computer Society, Los Alamitos, CA, USA. Pp165-181.

Sobalvarro P.G. and Weihl W.E. 1995. "Demand-based Coscheduling of Parallel Jobs on Multiprogrammed Multiprocessors". In *Job Scheduling Strategies for Parallel Processing*, *Lecture Notes in Computer Science*, Springer-Verlang, Berlin, Germany. Vol. 949. Pp106-126.

Squillante M.S., Wang F. and Papaefthymioy M. 1996. "Stochastic Analysis of Gang Scheduling in Parallel and Distributed Systems", Performance Evaluation, Elsevier, Amsterdam, Holland, Vol. 27&28 (4). Pp273-296.

Topcuoglu H., Hariri S. and Wu M-Y. 1999. "Task Scheduling Algorithms for Heterogeneous Processors". In *Proc. of the Eighth Heterogeneous Computing Workshop* (San Juan, Puerto Rico, April) IEEE Computer Society, Los Alamitos, CA, USA. Pp3-14.

Wang F., Papaefthymiou M. and Squillante M.S. 1997. "Performance Evaluation of Gang Scheduling for Parallel and Distributed Systems". In *Job Scheduling for Parallel Processing, Lecture Notes in Computer Science*, Springer-Verlang, Berlin, Germany. Vol. 1291. Pp184-195.

Zhang Y., Franke H., Moreira J. and Sivasubramaniam A. 2000. "The Impact of Migration on Parallel Job Scheduling for Distributed Systems". In *Proc. of Europar* (Munich, Germany, August) Europar, Passau, Germany. Pp242-251.

Zhang Y. and Sivasubramaniam A. 2001. "Scheduling Best-Effort and Real-Time Pipelined Applications on Time-Shared Clusters". In *Proc. of the 13th Annual ACM Symposium on Parallel Algorithms and Architectures* (Crete Island, Greece, July) ACM, New York, N.Y. Pp209-219.

BIOGRAPHY

HELEN D. KARATZA is an Associate Professor in the Department of Informatics at the Aristotle University of Thessaloniki, Greece. Her research interests include Performance Evaluation of Parallel and Distributed Systems, Multiprocessor Scheduling, Mobile Agents and Simulation. Helen Karatza is a member of the Editorial Board of the International Journal of Simulation: Systems, Science & Technology.