

Special issue on network-based high performance computing

H. Sarbazi-Azad · A. Shahrabi · H. Beigy

Published online: 19 March 2010
© Springer Science+Business Media, LLC 2010

Over the past decade, ever-increasing demands for greater computational power have necessitated the development of High Performance Computing (HPC) systems with high availability and reliability. This trend has transformed the traditional model of parallel processing into a model of computing where all components of HPC systems are applied together in a cooperative network in order to solve scientific problems of unprecedented complexity. The key element of HPC architectures is, of course, the underlying network since it aims to provide a low latency communication for parallelism. Network-based computing is now a subject of interest across the complete range of scales in which distributed systems operate, from those comprising multiple engines on a single chip to those harnessing the power of large numbers of powerful computers using wide area connections to implement grids or other cluster-based structures.

This special issue is devoted to presenting a range of relevant topics in the area of network-based HPC, covering a selection of its many aspects. We invited authors of selected papers from International CSI Computer Conference (CSICC2008) to

H. Sarbazi-Azad (✉) · H. Beigy
Department of Computer Engineering, Sharif University of Technology, Tehran, Iran
e-mail: azad@sharif.edu

H. Beigy
e-mail: beigy@sharif.edu

H. Sarbazi-Azad
School of Computer Science, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran
e-mail: azad@ipm.ir

A. Shahrabi
School of Computing and Engineering, Glasgow Caledonian University, Glasgow, UK
e-mail: a.shahrabi@gcal.ac.uk

extend their papers for consideration in the special issue. Also, we widely announced the special issue and invited researchers of the field to submit papers to the special issue. As a result, initially, 24 papers were submitted of which 18 papers were selected for review. After a rigorous review process, 13 papers were selected for publication in the special issue. What follows, immediately below, is a brief summary of these contributions intended to give a preliminary indication of the scope and variety of the issue.

Job scheduling has always been a fundamental issue as it has a direct impact on the performance of any multicomputer network. Focusing on both two-dimensional and three-dimensional mesh-connected multicomputers, a new non-FCFS job-scheduling scheme is proposed by Ababneh, Bani-Mohammad, and Ould-Khaoua. The proposed scheme aims to bound job waiting delays, while achieving superior performance in terms of higher system utilization and lower job turnaround times.

Peer-to-peer network has been recently appeared as any distributed network architecture composed of participants that make a portion of their resources available to other network participants without requiring any centralized coordination. The quality of live streaming in peer-to-peer multipath networks is the subject of the research conducted by Liu and Chen. Their study analyzes the key considerations and factors influencing live stream quality during system operations and improves present P2P (peer-to-peer) live streaming systems by allowing users to enjoy high quality of service under the limitations of network resources.

As another study under the peer-to-peer networks umbrella, Xhafa, Barolli, Caballe and Fernandez address the efficient management of peer groups in JXTA-based P2P systems as a key issue in many P2P applications that use peer group as a unit. Motivated by the need to support online teams of real virtual campuses, they propose the management of peer groups in JXTA-Overlay, a JXTA-based P2P middleware for the development of P2P applications.

The third paper in peer-to-peer networks area focuses on the effectiveness and scalability of search algorithms. By combining search and trust systems, Mashayekhi and Habibi propose a robust and efficient trust-based search framework for unstructured P2P networks. The proposed framework maintains limited size routing indexes, combining search and trust data to guide queries to most reputable nodes.

CONFIIT (Computation Over Network with Finite number of Independent and Irregular Tasks) is a purely decentralized peer-to-peer middleware for Grid computing, which has already been proposed. The main features and reaction of CONFIIT to topology changes is the subject of another study conducted by Flauzac, Krajecki and Steffenel. They demonstrate how the car-sequencing problem can be solved in a distributed environment to illustrate CONFIIT operation.

One of the more significant developments in network computing in recent times has been the emergence of Grid computing as a mechanism for harnessing processing resources and bringing them to bear on compute-intensive tasks. In such highly distributed environments, estimating the available bandwidth between clusters is a key issue for efficient task scheduling. In a research by Batista, Chaves, da Fonseca and Ziviani, the performance of two well-known available bandwidth estimation tools has been analyzed.

Distributed Real Time (DRT) systems are increasingly in demand of object profiling, scheduling and migration algorithms to respond to unpredictable transient

changes in load and availability of resources in an open environment. Du and Ruan propose a robust DRT model that does not require precise system parameters such as task execution times. By using the proposed model, it is easy to achieve coupling among processors and include various periodic and aperiodic tasks, load migration, and disturbance effects.

Load balancing for emerging and future highly distributed high-performance computing systems has always been an attractive research area and much progress has been reported relating to the design of new methods and algorithms. In another study conducted by Randles, Abu-Rahmeh, Johnson and Taleb-Bendiab, a scalable and reliable load-balancing scheme for the distributed resources accessible on Grid networks is demonstrated through matching the load on a resource network to approach regular connectivity on a network graph. The proposed scheme provides a distributed load balancing system by generating almost regular networks.

Scheduling parallel applications with precedence-constraints is emerging as a new challenge in volunteer computing systems. Choon Lee, Zomaya, and Siegel propose two novel robust task-scheduling heuristics, which identify best task-resource matches in terms of makespan and robustness. Both approaches are based on a proactive reallocation scheme enabling output schedules to tolerate a certain degree of performance degradation. Schedules are initially generated by focusing on their makespan. These schedules are scrutinized for possible rescheduling using additional volunteer computing resources to increase their robustness.

A striking feature of recent developments in networking has been the rise of wireless connectivity, but while this yields great benefits especially for roaming nodes, some issues such as network sustainability have also to be addressed. Focusing on ad hoc networks, Xu, Wang, and Srimani propose a self-stabilizing protocol for weakly connected dominating set in a given network graph. Self-stabilization is a protocol design paradigm that is especially useful in resource-constrained infrastructure-less networks since nodes can make moves based on local knowledge only and yet a global task is accomplished in a fault-tolerant manner.

Another research in the area of mobile ad hoc networks investigates the enhancement of routing protocols using the improvement achieved during broadcasting route requests in distance vector routing protocols. Yasin, Khalaf and Al-Dubai propose a new probabilistic method to improve the performance of existing on-demand routing protocols by reducing the route requests overhead during the route discovery operation.

Within the same networking environment, mobile ad hoc network, and in order to improve Media Access Control (MAC) under self-similar traffic, Abu-Tair, Min, Ni, and Liu propose an adaptive MAC scheme which dynamically adjusts the increasing function and resetting mechanism of contention window based on the status of network load. The performance of this scheme is investigated in comparison with the legacy of DCF under self-similar traffic and different mobility models.

Finally, Ysami and Mozaffari move the focus to anomaly detection algorithms used to obtain sufficient information about complex network traffic in intrusion detection systems. They propose, based on the k-means clustering and the ID3 decision tree learning approaches in machine learning theory, a combinatorial approach for unsupervised classification of anomalous and normal activities in computer network ARP traffic.

In closing, as guest co-editors, we express our thanks to the editor-in-chief of the Journal of Supercomputing, Professor H. Arabnia, for hosting this special issue devoted to network-based high performance computing and for his support and advice throughout the process of bringing the original conception to fruition. We also thank all the authors for their contributions, including those whose papers were not included in this special issue and, last, the many reviewers who contributed their time and energy to providing valuable evaluations and recommendations.