Introduction

Cees de Laat*, Chris Develder*, Admela Jukan*, and Joe Mambretti*

This topic is devoted to communication issues in scalable compute and storage systems, such as parallel computers, networks of workstations, and clusters. All aspects of communication in modern systems were solicited, including advances in the design, implementation, and evaluation of interconnection networks, network interfaces, system and storage area networks, on-chip interconnects, communication protocols, routing and communication algorithms, and communication aspects of parallel and distributed algorithms. In total 15 papers were submitted to this topic of which we selected the 7 strongest papers. We grouped the papers in two sessions of 3 papers each and one paper was selected for the best paper session. We noted a number of papers dealing with changing topologies, stability and forwarding convergence in source routing based cluster interconnect network architectures. We grouped these for the first session. The authors of the paper titled: "Implementing a Change Assimilation Mechanism for Source Routing Interconnects" propose a mechanism that can obtain the new topology, and compute and distribute a new set of fabric paths to the source routed network end points to minimize the impact on the forwarding service. The article entitled "Dependability Analysis of a Fault-tolerant Network Reconfiguration Strateg" reports on a case study analyzing the effects of network size, mean time to node failure, mean time to node repair, mean time to network repair and coverage of the failure when using a 2D mesh network with a fault-tolerant mechanism (similar to the one used in the BlueGene/L system), that is able to remove rows and/or columns in the presence of failures. The last paper in this session: "RecTOR: A New and Efficient Method for Dynamic Network Reconfiguration" presents a new dynamic reconfiguration method, that ensures deadlock-freedom during the reconfiguration without causing performance degradation such as increased latency or decreased throughput. The second session groups 3 papers presenting methods, protocols and architectures that enhance capacities in the Networks. The paper titled: "NIC-assisted Cache-Efficient Receive Stack for Message Passing over Ethernet" presents the addition of multiqueue support in the Open-MX receive stack so that all incoming packets for the same process are treated on the same core. It then introduces the idea of binding the target end process near its dedicated receive queue. In general this multiqueue receive stack performs better than the original single queue stack, especially on large communication patterns where multiple processes are involved and manual binding is difficult. The authors of: "A Multipath Fault-Tolerant Routing Method for High-Speed

* Topic Chairs.

H. Sips, D. Epema, and H.-X. Lin (Eds.): Euro-Par 2009, LNCS 5704, pp. 1013–1014, 2009.

Interconnection Networks" focus on the problem of fault tolerance for high-speed interconnection networks by designing a fault tolerant routing method. The goal was to solve a certain number of link and node failures, considering its impact, and occurrence probability. Their experiments show that their method allows applications to successfully finalize their execution in the presence of several faults, with an average performance value of 97% with respect to the fault-free scenarios. The paper: "Hardware implementation study of the Self-Clocked Fair Queuing Credit Aware (SCFQ-CA) and Deficit Round Robin Credit Aware (DRR-CA) scheduling algorithms" proposes specific implementations of the two schedulers taking into account the characteristics of current high-performance networks. A comparison is presented on the complexity of these two algorithms in terms of silicon area and computation delay. Finally we selected one paper for the special paper session: "A Case Study of Communication Optimizations on 3D Mesh Interconnects". In this paper the authors present topology aware mapping as a technique to optimize communication on 3-dimensional mesh interconnects and hence improve performance. Results are presented for OpenAtom on up to 16,384 processors of Blue Gene/L, 8,192 processors of Blue Gene/P and 2,048 processors of Cray XT3.