Enabling Energy Efficient Exascale Computing Applications with Optical Interconnects

John Shalf
Lawrence Berkeley National Laboratory, NERSC Division, 1 Cyclotron Road, Berkeley, California, 94720
Author e-mail address: jshalf@lbl.gov

Abstract: The past few years has seen a sea change in computer architecture that will impact every facet of our society as every electronic device from cell phone to supercomputer will need to confront parallelism of unprecedented scale. Enabling future advances in sustained computational performance will require fundamental advances in computer architecture and programming models that are nothing short of reinventing computing. Photonic interconnect technologies will play an essential role to enable future progress in exascale computing.
OCIS codes: (200.0200) Optics in Computing

1. Background

Computing technology has been a significant and pervasive driving force in the global technology market over the past two decades. It affects nearly every aspect of life from education, entertainment, transportation and personal communication to the basic infrastructure in our economy, medicine, engineering and science. Society has come to depend not just on computing, but on the increases in computing capability that have been available each year for given cost and power budget. However, for the first time in decades, the advances in computing technology are now threatened, because while transistor density is projected to increase with Moore's Law, the energy efficiency of silicon is not keeping pace. HPC system architectures are expected to change dramatically in the next decade as power and cooling constraints are limiting increases in microprocessor clock speeds. Consequently computer companies are dramatically increasing on-chip parallelism to improve performance. The traditional doubling of clock speeds every 18-24 months is being replaced by a doubling of cores or other parallelism mechanisms [1]. During the next decade the amount of parallelism on a single microprocessor will rival the number of nodes in the first massively parallel supercomputers that were built in the 1980s. Applications and algorithms will need to change and adapt as node architectures evolve. Future generation consumer electronics devices, which are limited by battery life, would not support any new features that rely on increased computing. The iPhone, Google, simulation-based medical procedures, and our understanding of climate change would not have been possible without these increases in computing performance [8]. If computing performance stalls at today's levels the Information Technology industry will shift from a growth industry to a replacement industry, and future societal impacts of computing will be limited to what can be done on today's machines.

The next major milestone in High Performance Computing, an exascale system, would be impractical at hundreds megawatts [2]. Computing technology is rapidly approaching a power wall, which will limit future growth in computing capability. Overcoming this power wall will require fundamental advances in component technologies using advanced nanomaterials to enable transformational changes in the power, performance, and programmability of future computing devices. The path towards realizing next-generation petascale and exascale computing is increasingly dependent on building supercomputers with unprecedented numbers of processors [3]. To prevent the interconnect from dominating the overall cost of these ultra-scale systems, there is a critical need for scalable interconnects that capture the communication requirements of ultrascale applications [4]. Future computing systems must rely on development of interconnect topologies that efficiently support the underlying applications' communication characteristics. It is therefore essential to understand high-end application communication characteristics across a broad spectrum of computational methods, and utilize that insight to tailor interconnect designs to the specific requirements of the underlying codes.

As scientific computing matures, the demands for computational resources are growing at a rapid rate. It is estimated that by the end of this decade, numerous grand-challenge applications will have computational requirements that are at least two orders of magnitude larger than current levels. However, as the pace of processor clock rate improvements continues to slow, the path towards realizing ultrascale computing is increasingly dependent on scaling up the number of processors to unprecedented levels. To prevent the interconnect architecture from dominating the overall cost of such systems, there is a critical need to effectively build and utilize network topology solutions with costs that scale linearly with system size.

OThH1.pdf

Among the many issues affecting scalability of future system is the scaling of high band- width interconnects, designed for both on-chip and off-chip communication to memory and to other computational devices. Future computing systems, whether based on traditional circuits or the proposal nanotechnology devices will rely on parallelism to keep power bud- gets manageable while increasing performance. At the macro-scale, interconnects must keep costs, efficiency, and power consumption under control in the face of exponential growth in system parallelism [5]. At the chip-level, nanophotonic interconnects [6,7] can exploit extremely high-capacity and low-power interconnection supported by inherent parallelism of optics. This document will describe requirements for future manycore processors with massively parallel nanophotonic and nanoelectronic interconnects for a new generation of logic elements. Contemporary computing systems do not have sufficient memory and communication performance to balance their computation rates primarily due to limited I/O throughput of the off-chip electrical links. Optics provide ultra-high throughput, minimal access latencies, and low power dissipation that remains independent of capacity and distance that would en- able I/O bandwidth to be uniformly plentiful across a system. Multi-wavelength operation can bring massive parallelism to the computing system to enable construction of systems that attack grand-challenge scientific problems such as the study of global climate change, and support continued growth in the data processing capabilities of commercial datacenters, which are estimated to double every 18 months. Massively parallel nanoelectronic interconnection offers low-power short distance interconnection between many cores.

This presentation will provide a deep analysis of the requirements of ultrascale applications in order to better understand the demands on hardware at the system scale [5] and chip scale [6,7] in the face of massive growth in parallelism. The application requirements for system-scale will be distilled into requirements for multi-tiered networks using the high-level abstraction of fit- trees. The next chapter will introduce the Hybrid Flexibly Assignable Switch Topology (HFAST) approach [5] to implementing a dynamically reconfigurable topology to enable a physical realization of the fit-tree method. Then we will revisit the application communication pattern analysis with an eye towards understanding chip-scale interconnect requirements for Networks on Chip (NoCs) that must offer scalable performance for interconnecting hundreds or even thousands of cores over the next decade. Finally, the performance study compares a number of different NoC topologies that include both electrical packet switch and optical circuit switch components [6,7].

[1] Krste Asanovic, Ras Bodik, Bryan Christopher Catanzaro, Joseph James Gebis, Parry Husbands, Kurt Keutzer, David A. Patterson, William Lester Plishker, John Shalf, Samuel Webb Williams, and Katherine A. Yelick. The landscape of parallel com- puting research: A view from Berkeley. Technical Report UCB/EECS-2006-183 (http://www.eecs.berkeley.edu/Pubs/TechRpts/2006/EECS-2006-183.html), EECS De-partment, University of California, Berkeley, December 2006

[2] DOE E3 Report: http://www.er.doe.gov/ascr/ProgramDocuments/ProgDocs.html

[3] DARPA Exascale Computing Study (TR-2008-13): http://www.cse.nd.edu/Reports/2008/TR-2008-13.pdf.

[3] IAA Interconnection Network Workshop, San Jose, California, July 21-22, 2008 http://www.csm.ornl.gov/workshops/IAA-IC-Workshop-08/

[4] Architectures and Technology for Extrame Scale Computing Workshop, San Diego, California, December 8-10, 2009, http://extremecomputing.labworks.org/hardware/index.stm

[5] S. Kamil, L. Oliker, A. Pinar, J. Shalf (2009). "Communication Requirements and Interconnect Optimization for High-End Scientific Applications." IEEE Transactions on Parallel and Distributed Systems, Volume . IEEE: 2009.

[6] Assaf Shacham, Keren Bergman, and Luca P. Carloni. Photonic networks-on-chip for fu- ture generations of chip multiprocessors. IEEE Transactions on Computers, 57(9):1246– 1260, 2008.

[7] G. Hendry, S.A. Kamil, A. Biberman, J. Chan, B.G. Lee, M Mohiyuddin, A. Jain, K. Bergman, L.P. Carloni, J. Kubiatocics, L. Oliker, J. Shalf (2009). **Analysis of Photonic Networks for Chip Multiprocessor Using Scientific Applications.** NOCS2009.

[8] John Shalf (2007). **" The New Landscape of Parallel Computer Architecture."** Journal of Physics: Conference Series, Volume . IOP Electronics Journals: 2007.