Advantages of Silicon Photonics for Multi-socket Systems

Advantages of Silicon Photonics for Multi-socket Systems

Scott Beamer, Krste Asanović
Department of Electrical Engineering and Computer Science
University of California, Berkeley, California
{sbeamer, krste}@eecs.berkeley.edu

Christopher Batten, Ajay Joshi, Vladimir Stojanović
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology, Cambridge, Massachusetts
{cbatten, joshi, vlada}@mit.edu

ABSTRACT

Future single-board multi-socket systems may be unable to deliver the needed memory bandwidth electrically due to power limitations, which will hurt their ability to drive performance improvements. Energy efficient off-chip silicon photonics could be used to deliver the needed bandwidth, and it could be extended on-chip to create a relatively flat network topology. That flat network may make it possible to implement the same number of cores with a greater number of small dies for a cost advantage with negligible performance degradation.

Categories and Subject Descriptors: B.4.3 [Computer Systems Organization]: Processor Architectures[Parallel Architectures]

General Terms: Design, Economics, Performance

Keywords: Silicon Photonics, Multi-socket

1. INTRODUCTION

Given the difficulties of scaling uniprocessor performance further, most commercial microprocessor manufacturers have instead used increased transistor densities to integrate multiple processor cores on one die [1]. To deliver further performance improvements, multi-socket systems have been used to increase the the computing power and memory capacity. These multi-socket systems will require increasing memory bandwidth to deliver realizable improvements in application performance. This bandwidth must come from not only connections to DRAM, but also from inter-socket links. Even if the bandwidth to these systems is not hampered by pin limitations, it will be restricted by power limitations from electrical off-chip signalling.

Silicon photonics could be used off-chip to solve this bandwidth problem, with its great potential for energy efficiency and bandwidth density. If photonics is used for the inter-socket links, it could also be extended on-chip closer to its destinations. In this work we present a scalable interconnect based on a monolithically integrated silicon photonics proposal, that is able to harness the technology's potential to create an uniform network topology. With an approximately flat multi-socket interconnect, the penalty for communicating between sockets is reduced, which may enable potential cost benefits from implementing the same aggregate die area over a greater number of smaller dies.

ICS'09, June 8–12, 2009, Yorktown Heights, New York, USA.
ACM 978-1-60558-498-0/09/06.

2. SILICON PHOTONICS POTENTIAL

Silicon photonics has emerged in recent years as an appealing way to enable high bandwidths without excessive area or power requirements [3, 5, 6]. Due to the diversity in prospective photonic technologies, we selected a particular monolithically integrated silicon photonic proposal [2, 4] to base our design and its evaluation on.

Since photonics uses light rather than electricity to transmit data, transmitted bits must undergo conversion at both ends (electro-optical and opto-electrical) which adds a constant latency and energy penalty. Because those penalties are constant, photonics excels over a distance due to greater amortization. The most compelling advantages for silicon photonics over forecasted electrical interconnects are its high bandwidth density and energy efficiency for off-chip communication. On-chip the selected technology performs well under the same metrics, but only if it travels a non-trivial distance. Using a coupler, it is possible to guide light from off-chip fibers onto on-chip waveguides without retransmission or modification. This enables seamless inter-chip links to be made, since if the constant conversion overhead is going to be paid for off-chip links, it makes sense to traverse the remaining on-chip portion optically as well since it is nearly free [2].

For the selected technology, laser light is generated in bulk off-chip, and carried by fiber to splitters on-chip where it will be directed to the various links. Power is consumed by photonic links at the endpoints on-chip for signaling as well as the off-chip light source. Along the path from the transmitter to the receiver, there are various types of losses the signal incurs, and sufficient laser power must be applied to compensate. The *optical critical path* is the path with the most loss from the light source to the last receiver, and it dictates how much laser power the system will need.

3. ARCHITECTURE

For this research we target single-board multi-socket systems, and our design leverages the potential of photonics to produce a flat network. These boards could be connected together by another network to create an even larger system, but within a board a core sees uniform memory performance. Since electrical interconnects are advantageous over short distances, we electrically join groups of cores (4–16) into *clusters* by shared L2 caches. These clusters are connected by dedicated photonic links to every memory controller (Figure 1). Fully connected networks are often avoided because of their quadratic growth in resource consumption, but the

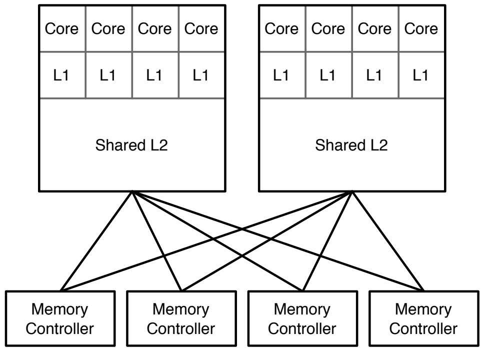

Figure 1: Topology for two clusters of four cores with four memory controllers

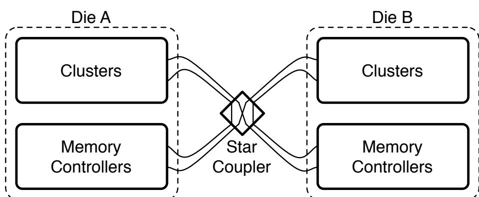

Figure 2: Logical view of a two die system

energy efficiency and the bandwidth density advantages of silicon photonics make it tolerable for a small system. A memory controller may actually communicate with multiple DRAM channels, but from the point of view of the network it is simply a point of arbitration for access to that memory. The links between the DRAM modules and the memory controllers are electrical because of the challenges involved with changing the DRAM interface, but future work could benefit greatly if these connections were photonic.

The simple network topology was not only chosen to make a flat network, but to also enable a single die design to be used in varying quantities to make a scalable range of systems. A cluster's memory bandwidth is uniformly spread across all the memory controllers, so in the maximum supported system size there is one direct channel between each cluster and each memory controller. For systems with less populated sockets, each cluster will get the same total bandwidth, but it will have multiple channels to each memory controller. To enable this flexibility, the actual connections between memory controllers and clusters is done off chip (Figure 2) so the changes necessary for systems of different sizes are localized to small off-chip components. To simplify the packaging and assembly of all the point-to-point connections, off-chip fibers are grouped into ribbons, which connect to a *star coupler*. The star coupler is a passive device that connects two groups of ribbons such that each ribbon has at least one fiber directly coupled with a fiber from every ribbon in the other group. Our design template is general enough that it is able to scale down to smaller dies while maintaining the same topology and nearly identical performance.

4. INCENTIVES FOR DISINTEGRATION

Using a greater number of smaller dies to implement the same silicon area could have cost advantages. Smaller dies should benefit from higher yield rates and increased tolerance to process variation, since they could be binned on finer granularities. A single reusable design will also have a higher sales volume, which will reduce non-recurring engineering (NRE) costs. This disintegration is made worthwhile by photonics, because otherwise it will increase the number of electrical pins and power spent on the interconnect. For our design, smaller dies will allow the system to be more spread out, which will reduce the power density and make it easier to electrically attach DRAM. Fixed costs per die (testing, packaging, and assembly) will cause penalties for using dies that are too small, but the optimum die size for cost may be smaller than current commercial designs.

5. RESULTS

Using our candidate technology, we evaluated the general design while varying the die size (16–256 cores/die) and the maximum supported system size (64–1024 cores). To scale to higher core counts will require a multi-hop network. The layout of each design was optimized to reduce the optical critical path loss because laser power can be the majority power consumer of a photonic interconnect. The area taken by the on-chip interconnect was always less than 10%, and the latency stayed roughly constant since the network topology stayed the same. Interestingly for the range of designs explored, independent of the total numbers of cores, systems with a modest number of dies (4–8) had the lowest optical power.

6. CONCLUSION

Silicon photonics provides an appealing way to supply the bandwidth needed to drive multi-socket systems, and a range of scalable designs capable of supporting up to 1024 cores with uniform memory bandwidth was presented. In a relatively flat network like the one presented, silicon photonics sufficiently reduces the barrier to going off-chip such that future die sizes may be chosen by what is most cost efficient instead rather what is most reasonable to manufacture.

7. REFERENCES

[1] K. Asanovic et al. The landscape of parallel computing research: A view from berkeley. Technical report, U.C. Berkeley, 2006.

[2] C. Batten et al. Building manycore processor-to-dram networks with monolithic silicon photonics. *High Performance Interconnects*, Jan 2008.

[3] N. Kirman et al. Leveraging optical technology in future bus-based chip multiprocessors. *IEEE Micro*, 27(6), Jan 2006.

[4] J. Orcutt et al. Demonstration of an electronic photonic integrated circuit in a commercial scaled bulk cmos process. *Conf. on Lasers and Electro-Optics*, 2008.

[5] A. Shacham et al. Photonic noc for dma communications in chip multiprocessors. *IEEE Symposium High-Performance Interconnects*, 15, Jan 2007.

[6] D. Vantrease et al. Corona: System implications of emerging nanophotonic technology. *ISCA*, Jan 2008.