Building Manycore Processor-to-DRAM Networks with Monolithic Silicon Photonics

Christopher Batten¹, Ajay Joshi¹, Jason Orcutt¹, Anatoly Khilo¹
Benjamin Moss¹, Charles Holzwarth¹, Miloš Popović¹, Hanqing Li¹
Henry Smith¹, Judy Hoyt¹, Franz Kärtnner¹, Rajeev Ram¹
Vladimir Stojanović¹, Krste Asanović²

¹ Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology, Cambridge, MA

² Department of Electrical Engineering and Computer Science
University of California, Berkeley, CA

Symposium on High Performance Interconnects
August 27, 2008
The manycore memory bandwidth challenge
The manycore memory bandwidth challenge

- 256 Cores
  - 4-way SIMD FMACs @ 2.5–5 GHz
  - 5–10 TFlops on one chip
  - Some apps require 1 byte/flop
  - Need 5–10 TB/s of off-chip I/O

Manycore Era

Year: 1980 to 2020

Number of cores
Cost of electrical processor-to-DRAM networks

256 Cores

- 4-way SIMD FMACs @ 2.5–5 GHz
- 5–10 TFlops and 5–10 TB/s
Cost of electrical processor-to-DRAM networks

256 Cores
- 4-way SIMD FMACs @ 2.5–5 GHz
- 5–10 TFlops and 5–10 TB/s

Off-chip I/O
- 5 pJ/b @ 10 Gb/s = 50 mW/b
- 4k–8k differential pin pairs
- 200–400 W
Cost of electrical processor-to-DRAM networks

256 Cores
- 4-way SIMD FMACs @ 2.5–5 GHz
- 5–10 TFlops and 5–10 TB/s

Off-chip I/O
- 5 pJ/b @ 10 Gb/s = 50 mW/b
- 4k–8k differential pin pairs
- 200–400 W

On-chip Interconnect
- 0.5–1 pJ/b @ 5 Gb/s = 2.5–5 mW/b
- 4k–8k bisection wires
- 10–40 W (just in wires)
Cost of electrical processor-to-DRAM networks

256 Cores
- 4-way SIMD FMACs @ 2.5–5 GHz
- 5–10 TFlops and 5–10 TB/s

Off-chip I/O
- 5 pJ/b @ 10 Gb/s = 50 mW/b
- 4k–8k differential pin pairs
- 200–400 W

On-chip Interconnect
- 0.5–1 pJ/b @ 5 Gb/s = 2.5–5 mW/b
- 4k–8k bisection wires
- 10–40 W (just in wires)

Can we use silicon photonics to help address the manycore memory bandwidth challenge?
Motivation

Photonic Technology

Network Architecture

Full System Design
Seamless On-Chip/Off-Chip Photonic Link

External Laser Source → Chip A
  Waveguide → Coupler → Transmitter
  → Single Mode Fiber → Chip B
  → Receiver

Chip A → Ring Modulator
Chip B → Photodetector, Ring Filter
Seamless On-Chip/Off-Chip Photonic Link

- Light coupled into waveguide on chip A
Seamless On-Chip/Off-Chip Photonic Link

- Light coupled into waveguide on chip A
- Transmitter off: Light extracted by ring modulator
Seamless On-Chip/Off-Chip Photonic Link

- Light coupled into waveguide on chip A
- Transmitter off: Light extracted by ring modulator
- Transmitter on: Light passes by ring modulator
Seamless On-Chip/Off-Chip Photonic Link

- Light coupled into waveguide on chip A
- Transmitter off: Light extracted by ring modulator
- Transmitter on: Light passes by ring modulator
- Light continues to receiver on chip B
Seamless On-Chip/Off-Chip Photonic Link

- Light coupled into waveguide on chip A
- Transmitter off: Light extracted by ring modulator
- Transmitter on: Light passes by ring modulator
- Light continues to receiver on chip B
- Light extracted by receiver’s ring filter and guided to photodetector
Photonic Component Characterization

**Standard CMOS process**

- Waveguides
- Ring Modulators
- Ring Filters
- Photodetectors

Simulation

65 nm Test Chip
Photonic Component: **Waveguide**

- **Motivation Photonic Technology Network Architecture – Full System Design**

**Photonic Component: Waveguide**

- External Laser Source
- Chip A:
  - Waveguide
  - Coupler
  - Ring Modulator
- Chip B:
  - Photodetector
  - Ring Filter

- Polysilicon waveguides • Etched air gap for cladding
  - Target 4 µm pitch increases bandwidth density

- Backend Dielectric
- Polysilicon Transistor Gates
- Etch Hole
- Polysilicon Waveguides
- Silicon Substrate
- Shallow Trench Isolation
- Air Gap
Photonic Component: **Ring Modulator**

Small 10 μm diameter rings and monolithic integration decrease parasitics.

Estimated energy: \(<100 \text{ fJ/b} \) (circuits) + 100 fJ/b (thermal tuning)

Estimated data rate: 10 Gb/s
Photonic Component: **Ring Filter**

Cascaded double ring design improves frequency selectivity
Estimated number of wavelengths per waveguide: **64**
Can send wavelengths in opposite directions down same waveguide
Photonic Component: **Photodetector**

- Embedded SiGe used to detect ~1200 nm light
- Monolithic integration enables waveguide to be close to detector for good optical coupling
- Sub-100 fJ/b receiver energy seems feasible
- Still work to be done on detector sensitivity
Silicon photonic’s energy and area advantage

<table>
<thead>
<tr>
<th>Link Type</th>
<th>Energy (pJ/b)</th>
<th>Bandwidth Density (Gb/s/μm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global on-chip photonic link</td>
<td>0.25</td>
<td>160-320</td>
</tr>
<tr>
<td>Global on-chip optimally repeated M9 wire in 32 nm</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Off-chip photonic link (50 μm coupler pitch)</td>
<td>0.25</td>
<td>13-26</td>
</tr>
<tr>
<td>Off-chip electrical SERDES (50 μm pitch)</td>
<td>5</td>
<td>0.2</td>
</tr>
<tr>
<td>On-chip/off-chip seamless photonic link</td>
<td>0.25</td>
<td></td>
</tr>
</tbody>
</table>
Motivation

Photonic Technology

Network Architecture

Full System Design
Leveraging silicon photonics to address the memory bandwidth challenge

![Diagram showing the architecture with leveraged silicon photonics](image-url)
Baseline Network Architecture: Mesh Topology

Logical View

Physical View

Request Network

Response Network

DRAM Module

DRAM Module

DRAM Module

DRAM Module

Mesh Router

Router and Access Point

DRAM Module

DRAM Module

DRAM Module

DRAM Module
Analytical modeling of energy and throughput tradeoffs

- 22 nm – 256 cores @ 2.5 GHz
- Performance will most likely be energy constrained
- Fixed 8 nJ/cycle energy budget (20W)
- Use simple gate-level models to estimate energy, ideal throughput under uniform random traffic, and zero-load latency
Analytical modeling of energy and throughput tradeoffs

- 22 nm – 256 cores @ 2.5 GHz
- Performance will most likely be energy constrained
- Fixed 8 nJ/cycle energy budget (20W)
- Use simple gate-level models to estimate energy, ideal throughput under uniform random traffic, and zero-load latency
Ideal throughput vs. off-chip I/O energy efficiency

- Decreased off-chip I/O energy, results in more I/O bandwidth and mesh bandwidth
- Latency decreases slightly due to lower serialization latency
- In photonic range almost all of the energy is being spent on the mesh
- A more energy efficient on-chip interconnect should further improve throughput
Mesh Augmented with Global Crossbar

**Logical View**

- **Group A**
  - Request Network A
  - Switch
  - DRAM Module
  - Switch
  - Response Network A
- **Group B**
  - Request Network B
  - Switch
  - DRAM Module
  - Switch
  - Response Network B

**Physical View**

- **DRAM Module**
  - Switch
  - DRAM Module
Analytical modeling of global crossbar topology

- Global crossbar increases energy efficiency of the on-chip interconnect improving throughput
- Global traffic is moved from energy-inefficient mesh channels to energy-efficient on-chip silicon photonics
- Global crossbar has little impact in the electrical range since very little energy is being spent in the on-chip interconnect to begin with
- Latency decreases due to lower serialization and hop latency
Simulation Methodology

- Execution driven cycle-accurate network simulator
- Models pipeline latencies, router contention, credit-based flow control, and serialization overheads
- Configuration same as in analytical modeling except:
  - Mesh networks use dimension ordered routing
  - 16 DRAM modules distributed around chip
  - Memory channels cache-line interleaved
  - Normalized buffering in terms of bits
Simulation Results

- Synthetic uniform random traffic with 256 bit messages
- For simple mesh (no groups) we see a $\approx 2 \times$ improvement in throughput at similar latency
Simulation Results

- Synthetic uniform random traffic with 256 bit messages
- For simple mesh (no groups) we see a $\approx 2 \times$ improvement in throughput at similar latency
- Adding global crossbar improves performance of photonic system but has little impact on electrical system
- Throughput is improved by $\approx 8-10 \times$ and best throughput is $\approx 5$ TB/s
Motivation

Photonic Technology

Network Architecture

Full System Design

Motivation

Photonic Technology

Network Architecture

Full System Design
Simplified 16-core system design
Simplified 16-core system design
Simplified 16-core system design
Simplified 16-core system design
Simplified 16-core system design
Full 256-core system design

Estimated area for photonics: 5-10%
Estimated total laser power: 6.5 W
Advantages of photonics for packaging and system-level integration
Advantages of photonics for packaging and system-level integration
Take Away Points

- Silicon photonics is a promising technology for increasing the energy efficiency and the bandwidth density for on-chip and off-chip interconnect.

- Addressing the manycore bandwidth challenge requires implementing both global on-chip interconnect and off-chip I/O with photonics.

- We can efficiently implement global all-to-all connectivity with silicon photonics by using vertical waveguides, horizontal waveguides, and a ring filter matrix where they cross.