Si-Photonics Technology
Towards fJoule/bit Optical Communication in Many-core Chips

Adaptive Systems Lab
Abderazek Ben Abdallah
benab@u-aizu.ac.jp
Contents

1. Trends in CPU
2. Optical interconnect
3. Si-Photonics Many-core chips
4. Research direction, challenges
5. Concluding remarks

December 21, 2014

ASCI Blue-Mountain (1.6 TeraOps, 929 m², 1.6 Mwatts)

Ref. http://www.jipdec.or.jp

Intel's Tera-scale 80 core Chip
(1.63 Teraflops @ 5.1 GHz, 175 watts, and 1.81 Teraflops @ 5.7 GHz, 265 watts).

ASCI Blue-Mountain: 1.6 TeraOps, 929 m², 1.6 Mwatts

Ref. http://www.jipdec.or.jp

Sonny Play station 3 (2006): 1.8 Teraflops, 0.08 square meter, <200 watts
Moore’s Law

Current Processor Research Trends

Intel Knights Corner 50 cores, 200 Threads

Oracle T5 16 cores, 128 Threads

IBM Power 7 8 cores, 32 threads

Nvidia Fermi 540 CUDA cores

Intel 4004 (1971): 4-bit processor, 2312 transistors, ~100 KIPS, 11 mm² chip

1000s of processor cores per die could be integrated? → How about power scaling!
Wire and I/O scaling problems

Energy cost of data movement relative to the cost of a flop for current and 2018 systems. (Shalf et al., VECPAR 2010)

• Preparing the operands costs more than performing computing on them!
• There is no Moore’s law for communications!
Current Processor Research Trends

- Easier Programming
- Easier Implementation
- Low energy efficiency
- No specific HW for different tasks.

• Conventional Electric-wiring on chip (add-hoc wiring) consumes half of CPU power.
• Teraflop Chip router consumes 28% of CPU power.
Limitations of Traditional E-NoC

- Multi-hop communication.
- Receive, buffer and retransmit every bit at every switch.
- High latency and energy dissipation especially in large system.
Limitation of Electric/Metal wire

- Electronics is not good at high bit/s communication.

Can we have photonic networking?
Energy Cost for Communication

Small transmission energy, but high processing energy.

Conventional photonics

Required energy cost


Miller (IEEE Proc. 2009)

fJ/bit
We need a “Spring-Revolution” to deal with the Power/Energy Wall!

• The computation power of CPU is still progressing exponentially, and there are strong demands to keep this progress rate for the next decades.

• If we assume the same progress rate, the allowable energy for transmitting a single bit in a chip should be around a few fJ in 2025 [Miller 2009].

• The problem of on-chip electric communication is largely attributed to the finite RC of wirings.
  – As the bit rate goes up, we have to use wider and shorter wires in order to avoid the RC delay → Conflicting with the limited space-budget in a chip!!.
1. Trends in CPU

2. Optical interconnect

3. Si-Photonics Many-core chips

4. Research direction, challenges

5. Concluding remarks
Why optical interconnection?

- Larger bandwidth is possible for a long wire
  - Bandwidth can be enhanced by WDM.
- Efficient Energy at high bit rate communications.
  - No energy cost for transfer (no charging energy)
- A photon can generate ≈1 volt (via photo-electric effect), which is NOT bound by the light intensity (number of photons).
Transmission over Si Wire/Waveguide

Snell’s Law of Refraction:

\[
\frac{\sin \theta_1}{\sin \theta_2} = \frac{n_2}{n_1} = \frac{v_1}{v_2}
\]
Total internal reflection in Si Wire/Waveguide

Let $\theta_2 = \pi/2$:

Then $\sin \theta_1 = \frac{n_2}{n_1}$

$\theta_c = \sin^{-1} \left( \frac{n_2}{n_1} \right)$

For $\theta_1 > \theta_c$, light ray is completely reflected.

---

December 21, 2014

benab@u-aizu.ac.jp
Total internal reflection in Si Wire/Waveguide

Total internal reflection keeps all optical energy within the core, even if the fiber bends.

$\theta_1 < \theta_2$

$n_{\text{cladding}} < n_{\text{core}}$
Main components

- **Laser Source**: Inject the required laser lights into waveguide
- **Modulators**: Modulate the laser lights to ‘0’ and ‘1’ states
- **Photodetectors**: Detect the laser lights and convert to electrical signal
- **Turn Resonators**: Control the routing direction of the laser lights
Problems in Photonic Integration

• Fabrication cost → Being explored by Si photonics.

• Low energy cost for data transmission
  → This is a big issue. How much should we reduce?

• Larger scale with higher density → What applications for large-scale photonics?
Contents

1. Trends in CPU
2. Optical interconnect
3. Si-Photonics Many-core chips
4. Research direction, challenges
5. Concluding remarks
Typical Packet format

<table>
<thead>
<tr>
<th>Tail flit</th>
<th>Body flit</th>
<th>Head flit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ending</td>
<td>Carried</td>
<td>Flit</td>
</tr>
<tr>
<td>flit</td>
<td>Payload</td>
<td>information</td>
</tr>
</tbody>
</table>

OASIS-1: Overview of Electronic Packet Switched NoC

Scalability issue if chip is very large → Latency, bandwidth, and power problems.

Multihop communication
Receive -> Buffer -> Transmit every flit at every switch.

R: Router. NI: Network interface. PE: Processing Element
OASIS-1: Overview of Electronic Packet Switched NoC

Router addressed NM (in decimal)

Wire length reduction

Footprint reduction

3D- Network-on-Chip architecture

Lateral link (1mm ~ 4mm)
Vertical link (10 μm ~ 200 μm)
OASIS-1: Overview of Electronic Packet Switched NoC

Power: 222.387 uW, Number of Pins: 557

OASIS Network-on-Chip System

Credit: Y. Matsumoto
OASIS-1: Overview of Electronic Packet Switched NoC

Table 1: Simulation configuration.

<table>
<thead>
<tr>
<th>Parameters / System</th>
<th>LAFT</th>
<th>LA-XYZ</th>
<th>XYZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network Size</td>
<td>Matrix</td>
<td>3x6x6</td>
<td>3x6x6</td>
</tr>
<tr>
<td>Mesh</td>
<td>Transpose &amp; Uniform</td>
<td>4x4x4</td>
<td>4x4x4</td>
</tr>
<tr>
<td>Bit size</td>
<td>34 bits</td>
<td>34 bits</td>
<td>31 bits</td>
</tr>
<tr>
<td>Header size</td>
<td>13 bits</td>
<td>13 bits</td>
<td>10 bits</td>
</tr>
<tr>
<td>Payload size</td>
<td>21 bits</td>
<td>21 bits</td>
<td>21 bits</td>
</tr>
<tr>
<td>Buffer Depth</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Switching</td>
<td>Wormhole-like</td>
<td>Wormhole-like</td>
<td>Wormhole-like</td>
</tr>
<tr>
<td>Flow control</td>
<td>Stall-Go</td>
<td>Stall-Go</td>
<td>Stall-Go</td>
</tr>
<tr>
<td>Scheduling</td>
<td>Matrix-Arbiter</td>
<td>Matrix-Arbiter</td>
<td>Matrix-Arbiter</td>
</tr>
<tr>
<td>Routing</td>
<td>Look Ahead Fault Tolerant</td>
<td>Look-Ahead-XYZ</td>
<td>XYZ</td>
</tr>
</tbody>
</table>

Table 3: Hardware complexity comparison results.

<table>
<thead>
<tr>
<th>Target device</th>
<th>System</th>
<th>Area</th>
<th>Power (mW)</th>
<th>Speed (MHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Static</td>
<td>Dynamic</td>
</tr>
<tr>
<td>FPGA</td>
<td>LAFT</td>
<td>3272</td>
<td>1296.92</td>
<td>178.09</td>
</tr>
<tr>
<td></td>
<td>LA-XYZ</td>
<td>3093</td>
<td>1264.36</td>
<td>169</td>
</tr>
<tr>
<td></td>
<td>XYZ</td>
<td>2809</td>
<td>1258.01</td>
<td>165</td>
</tr>
<tr>
<td>Structured-ASIC</td>
<td>LAFT</td>
<td>32420</td>
<td>281.6</td>
<td>17.1</td>
</tr>
<tr>
<td></td>
<td>LA-XYZ</td>
<td>30646</td>
<td>274.25</td>
<td>16.24</td>
</tr>
<tr>
<td></td>
<td>XYZ</td>
<td>27822</td>
<td>277.6</td>
<td>15.83</td>
</tr>
</tbody>
</table>

Figure 10: Latency per flit evaluation with: (a) Transpose (b) Uniform (c) 6x6 Matrix.

Figure 11: Throughput evaluation with: (a) Transpose (b) Uniform (c) 6x6 Matrix.
PHENIC: Hybrid Si-Photonic NoC
Replace Wires with Waveguides and Electrons with Photons!

**Electrical NoC**
- Buffer, receive and re-transmit at every switch
- Off chip is pin-limited
- Large power/energy

**Electrical-Photonic NoC**
- Modulate/receive ultra-high bandwidth data stream once per communication.
- Switch routes entire multi-wavelength high BW stream
- Low power switch fabric, scalable
Basic Optical Switching Element

1. crossing element

2. parallel element
PHENIC: Hybrid Si-Photonic NoC
Replace Wires with Waveguides and Electrons with Photons!
PHENIC: Hybrid Si-Photonic NoC
Replace Wires with Waveguides and Electrons with Photons!
Routing in Hybrid Si-Photonic NoC

PHENIC: Hybrid Si-Photonic NoC
Replace Wires with Waveguides and Electrons with Photons!
Routing in Hybrid Si-Photonic NoC

1. Reserve the path

2. ACK

3. Transmit data on the Photonic layer

4. Release (tear-down)
1. Reserve the path
   - A path setup message is sent by the source in the electrical network to establish a path for the optical network.

2. ACK
   - A pulse is sent back to the source node by the destination node in the optical network, and optical data can be transferred.

3. Transmit data on the Photonic layer

4. Release (tear-down)
   - Teardown message is sent by the source node in the electrical control network to release the optical circuit.
PHENIC: Hybrid Si-Photonic NoC
Replace Wires with Waveguides and Electrons with Photons!

E-Router for Path Setting and Short Messages
PHENIC: Hybrid Si-Photonic NoC

Replace Wires with Waveguides and Electrons with Photons!

Bandwidth, power and latency

- Bandwidth (Gb/s)
- Latency (ns)
- Power (W)

Network sizes: 16x16, 8x8, 4x4

Benchmarks: Random, neighbor, Bitreverse

Networks: PNoC-Mesh, ENoC-Mesh
Contents

1. Trends in CPU
2. Optical interconnect
3. Si-Photonics Many-core chips
4. Research direction, challenges
5. Concluding remarks
Si-Photonics in computing system today

- Optical I/O’s for chip-to-chip and chip-to-board links (IBM, Intel, Fujitsu)
- E-O-E transceivers for Opto-Silicon Interposer
Optical link

- Uses monolithic integration that reduces energy consumption
- Utilizes the standard bulk CMOS flow
- Cladding is used to increase the total internal reflection $\Rightarrow$ reduces data loss
Photonics in computing system

Transmission over fiber (WDM)

$\lambda_1 \lambda_2 \lambda_3 \ldots \lambda_n$

channel

$>1 \text{ TBps}$

$<1 \text{ mW/Gbps}$

DRAM

Receiver/Transmitter

Multicore Processor (CMP)

WDM, DWDM

- Supports WDM that improves bandwidth density
- DWDM can transports tens to hundreds of wavelengths per fiber.
- Integrated Tb/s optical link on a single chip is ongoing
Current Research in Photonic Components

A reversely biased p-i-n diode to eliminate the TPA-induced FCA

Simulated Raman Scattering (SRS)

IBM/Columbia Si Wire/Waveguide

Si Wire/Waveguide

Laser

Modulator

Photodetectors

December 21, 2014

benab@u-aizu.ac.jp
Photonic Components and Future Demands

- The necessity of low energy in optical output devices, with a ~ 10 fJ/bit device energy target emerging.
  - Some Modulators and lasers meet this requirement
  - Low (few fF or less) photodetector capacitance is important
  - Very compact wavelength splitters are essential
  - Dense waveguides are also necessary on chip or on-boards for guided wave optical scheme.
Contents

1. Trends in CPU
2. Optical Interconnect
3. Si-Photonics Many-core Chips
4. Current Research direction
5. Concluding remarks
Concluding remarks

• Nanophotonics will play a crucial role for on-chip interconnects

• Several technologies:
  – Si photonics, high-index contrast waveguides, photonic crystals, and plasmonics.

• Si-Photonics design approach can reduce total energy, and improve system throughput by 15-20x
  – Several approaches have been explored
  – Much more other studies should be done
References

References

A. Ben Abdallah

Multicore Systems On-Chip: Practical Software/Hardware Design

Series: Atlantis Ambient and Pervasive Intelligence, Vol. 7

- Provides practical hardware/software design techniques for Multicore Systems-on-Chip
- Provides a real case study in Multicore Systems-on-Chip design
- Provides interaction between the software and hardware in Multicore Systems-on-Chip
- Provides detailed overview of various existing Multicore SoCs

2013, XXVI, 273 p. 196 illus., 79 illus. in

http://aslweb.u-aizu.ac.jp/
Thank you!

Abderazek Ben Abdallah
benab@u-aizu.ac.jp

The University of Aizu
To Advance Knowledge for Humanity