Ultraperformance Nanophotonic Intrachip Communications: UNÍC

Jagdeep Shah
DARPA/MTO

Frontiers of Extreme Computing 2007
Santa Cruz, CA
October 21-24, 2007

Approved for Public Release, Distribution Unlimited
UNÍC: Outline

• Introduction: Vision and Challenges
  • Microprocessor Challenges
    • Photonics/EPIC
      • UNÍC
    • Summary
Device Scaling

Microprocessor Scaling

Intel 4004 (1971)
2312 transistors, 11 mm²
~20,000 transistors/cm²

Intel Itanium 2 (2006)
1.72 billion transistors, 596 mm²
~300,000,000 transistors/cm²

15,000 X Increase in Device Density

The Impact of Moore’s Law:
• Device scaling most obvious in microprocessors
• On the path to >> 1 billion devices/cm²
• 3D layer stacking on all foundry roadmaps

Microprocessors are becoming Ultradense Systems

DARPA/MTO VISA Program
Other Ultradense systems

DARPA MoleApps – Aim: $10^{15}$ devices/ cm$^3$

17 nm half-pitch, $3.5 \times 10^{11}$ /cm$^2$ demonstrated

Courtesy HP
CHALLENGE

• How will these ultradense functional units communicate
  – With each other?
  – With the external world?

VISION

DEVELOP A PATHWAY FOR SUCH COMMUNICATIONS
• DARPA is about to launch a program to develop such a pathway:
  • for communications between ultrahigh functional units on a chip and from the chip to the outside world

• Enormous challenges: Focus on a specific challenge of microprocessors

• Use microprocessor circa 2017 as a design driver
MICROPROCESSOR CHALLENGES
We’re already seeing this effect: Microprocessor performance is not keeping pace with device scaling due to limitations at the circuit level

- Diminishing returns of ILP
- Thermal constraints
- Increasing complexity
- Growing communications gap

System-level balance
- Memory/ Bisection-Bandwidth balanced against peak performance
  Metric: Bytes/FLOP ([B/s]/[FLOPS]) ~1
- On-chip/ off-chip bandwidth balanced
- Power Consumption
  Uniform Distribution: 1/3 Processing, 1/3 Communications, 1/3 Memory

Communications challenges prevent actual system performance from meeting theoretical peak performance

Supercomputer Comms Gap
- Memory and bisection bandwidth imbalance a growing problem
- Limited by power consumption

Example: Cray XT4, 380 TF Bytes/Flop ~0.06
Multicore architectures designed to deliver increased performance at the circuit level

1 TFLOPS on-chip reported

“Supercomputers on-Chip”
Multicore Processor Plane

Challenges

- Reducing system-level power consumption
- High-bandwidth, low-power access to ultradense devices
- The communications gap:
  - Latency/Bandwidth limits
  - Power dissipation as wires shrink
  - Power-hungry off-chip communications to memory

- Processor/On-chip Memory
  - 100-1000 compute cores in ~6 cm²
  - ~10 TFLOPS peak performance
  - 3D-integration
  - >100 billion active devices!

- Required Communications Network
  - ~80 Tb/s on-chip bandwidth
  - ~80 Tb/s off-chip memory BW

- Total System Power ~200 W

Electronics cannot meet communication requirements within power budget. Actual performance cannot meet theoretical peak performance.

A new communications strategy is required
The high-bandwidth, low-latency properties of optical interconnects are continually meeting our increasing communication needs at smaller scales.

*Increased Bandwidth and Device Density*

Today

In Development

Future

Board to Board

Chip-to-Chip

On Chip


Is photonic signaling a potential solution to the chip-scale communications problem?
### BENEFITS

- **High Bandwidth/Capacity**
  - Wavelength Division Multiplexing (WDM)
  - Time Division Multiplexing (TDM)
  - Space Division Multiplexing (SDM)
- **No need for power hungry repeaters**
- **Power Independent Of Distance (\(\sim 10\) cm)**
- **Seamless I/O**

### CHALLENGES

- **Compatibility with Silicon**
- **Current devices too large: can they be scaled down in size?**
- **Current devices are power hungry: can the power be reduced?**
- **Spectral bandwidth vs. thermal stability**

Devices required to meet these challenges will be very different from today’s devices performing the same functions.
Success of EPIC gives confidence that this is possible.
• Legacy microphotonic devices were discrete components using a variety of materials
• Moore’s law has opened the way for integrated photonics in Si
• High index-contrast of Si/SiO\textsubscript{2} and smooth features below 90nm node enable nanophotonic devices
• DARPA EPIC Program has demonstrated high performance, monolithically integrated photonic and electronic devices using a standard CMOS foundry
• Application-specific EPIC chips demonstrated
UNÍC

PHOTONICALLY-ENABLED MICROPROCESSOR
UNÍC APPROACH

- Select a Design Driver (with ultradense functional units: microprocessor circa 2017)
- Design a photonic communication network within the constraints imposed by the design driver
- Quantify the system benefits of the approach
- Quantify device requirements to enable such communications
- Demonstrate the required device performance
- Demonstrate on-chip functional communication links with all essential components working in unison (sufficiently aggressive to convince the skeptics)
- Demonstrate multiple high performance microprocessors communicating via on-chip optical links
UNÍC: OBJECTIVE

Demonstrate to the microprocessor community (and others) that photonic intra-chip – and seamless off-chip – communication is a credible technology that will allow actual system performance to scale to a level not possible with electronic communications.

Comprehensive, team-based efforts encompassing:
1. Photonic Communication Architecture (on/off chip)
2. Device demonstrations (compatible with CMOS fabs) far beyond EPIC
3. System-level performance-benefit analysis
4. Full-link demonstrations of all critical technologies working together; application emulation
5. Microprocessors communicating via on-chip optical links

No effort on processor design, architecture…
Optical Link Key Components

How can 80 Tb/s be achieved?
• Transmit – Spatial Division Multiplexing (SDM) and Wave Division Multiplexing (WDM) of optical signals with multiple, low-insertion-loss, high-BW modulators

• Route – Spectrally filtered passive networks or active, arbitrated spatial switch networks using low loss, high power handling waveguides. Power consumption due to tuning must be kept at a minimum by using thermally tolerant designs.

• Receive – High-BW, high-responsivity, low-power (no TIA) photo detectors

Link Energy ~100 fJ/bit
# Photonic Device Challenges

**System-level constraints mandate stringent performance requirements**

<table>
<thead>
<tr>
<th>Device Example (w/drivers)</th>
<th>EPIC Demonstrated</th>
<th>UNÍC Requirements</th>
<th>Required Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>10+ Gb/s Modulator</strong></td>
<td>Area 700 µm²</td>
<td>28 µm²</td>
<td>25 X</td>
</tr>
<tr>
<td></td>
<td>Power 330 mW</td>
<td>0.8 mW</td>
<td>825 X</td>
</tr>
<tr>
<td><strong>WDM Filter</strong></td>
<td>Area 0.6 mm²</td>
<td>0.005 mm²</td>
<td>120 X</td>
</tr>
<tr>
<td></td>
<td>Power 73 mW (tuning)</td>
<td>0 mW ?</td>
<td>???</td>
</tr>
<tr>
<td><strong>10+ Gb/s Detector</strong></td>
<td>Area 0.16 mm²</td>
<td>0.01 mm²</td>
<td>16X</td>
</tr>
<tr>
<td></td>
<td>Power 36 mW</td>
<td>1 mW</td>
<td>36 X</td>
</tr>
</tbody>
</table>

**UNÍC will consist of many EPIC-like circuits with dramatically reduced power consumption and dimensions**

**Next Generation Intra-chip Communication Devices Require Dramatic Size and Power Reductions**
Potential Benefits

PHOTONIC TECHNOLOGY BENEFITS
• High Bandwidth: Wavelength, Time, and Space Division Multiplexing
• Low Power: No repeaters, buffers, regenerators; independent of dist.
• Seamless I/O: No need for power-hungry off-chip communication
• Enables high performance for low system power

System-level benefits
– Restores B/F system balance to maximize actual performance
– Facilitates architectures which reduce programming complexity (e.g. shared memory)
– Reduces chip and system power
– Enables deployable, chip-scale supercomputers

Real-Time, High-Performance Embedded Processing

Supercomputers

Image Processing  SAR Processing  Autonomous Ops  Cryptanalysis

Frontiers of Extreme Computing 2007: Santa Cruz, CA
Approved for Public Release, Distribution Unlimited
SUMMARY

- Device scaling is producing ultradense electrical microsystems
- Microprocessors are becoming ultradense supercomputers on-chip
- High-performance computer systems are becoming unbalanced due to communications bottleneck
- Photonic communications may provide a novel solution to high BW on- and off-chip communications challenges
- UNIC addresses this problem head-on