

Monolithic Silicon Photonics

Vladimir Stojanović, Judy Hoyt, Rajeev Ram,
Franz Kaertner, Karl Berggren, Martin Schmidt,
Henry Smith, Erich Ippen, and **Krste Asanović***

Massachusetts Institute of Technology

*University of California at Berkeley

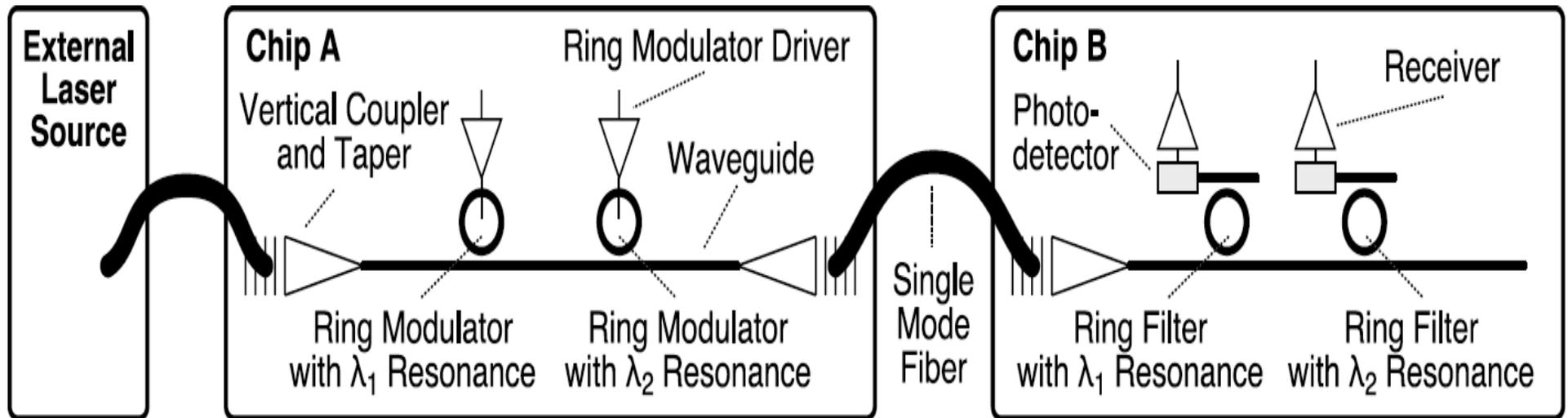


IAA Interconnection Network Workshop, July 22, 2008

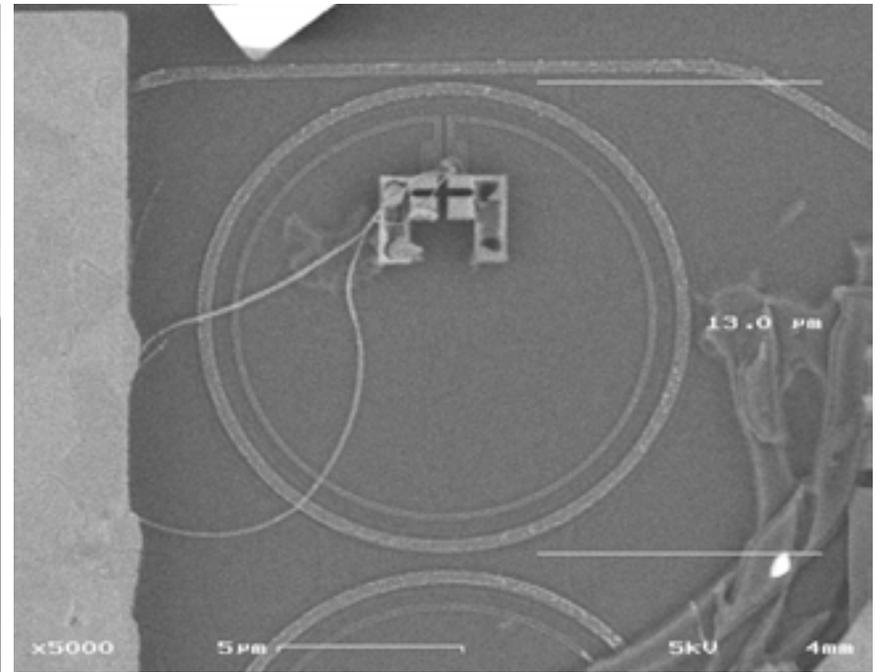
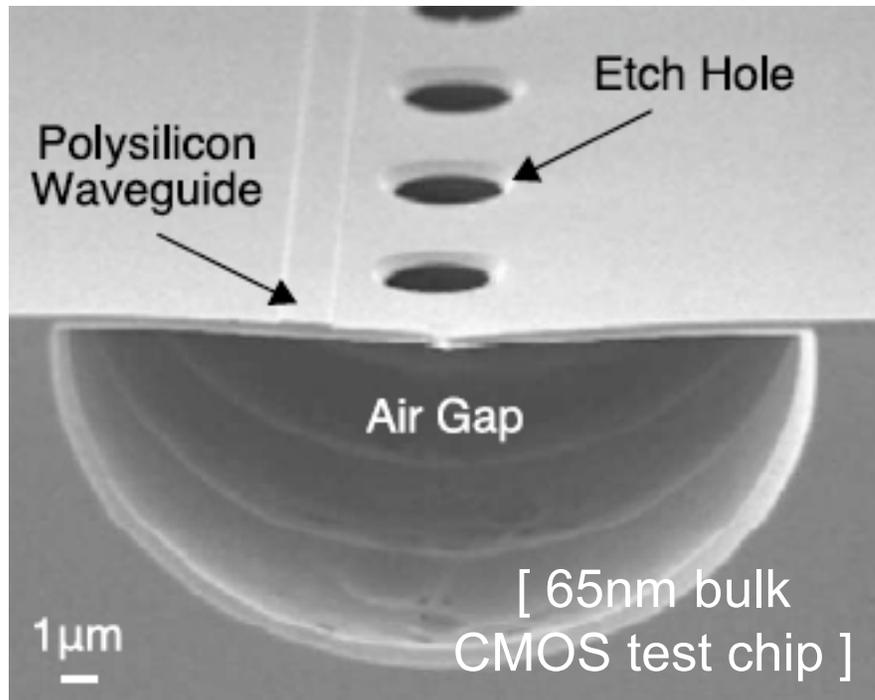
Outline

- ❑ Monolithic silicon photonic technology
 - Waveguides, modulators, and detectors in standard *bulk* CMOS flow. (Compatible with commodity logic technology and possibly DRAM technology).
 - Dense Wavelength-Division Multiplexing (DWDM) gives extremely high bandwidth density
- ❑ Manycore Processor-DRAM interconnect architecture based on opto-electronic crossbar
 - High-bandwidth, low-energy access to shared memory system within processor module
 - Around 10x improvement in bandwidth over optimistic electrical system for power-constrained system

Basic DWDM Photonic Link Path

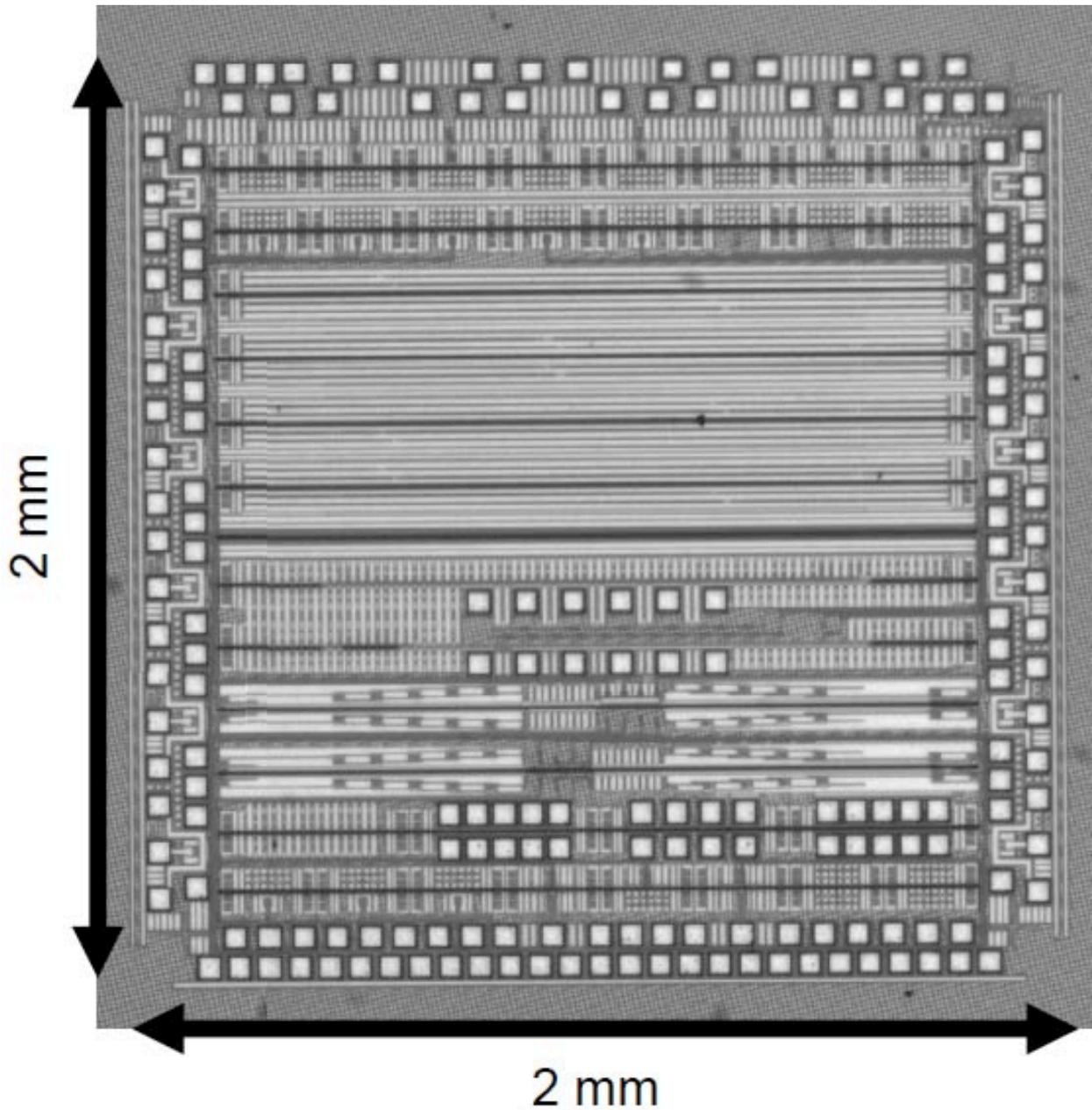


Waveguide technology



- ❑ Standard bulk CMOS processing followed by under-etch of waveguide areas through sacrificial vias (<math><10\text{dB/cm}</math> @1550nm)
 - Need around 5μm gap to provide effective cladding
 - Also provides thermal isolation for ring resonators
- ❑ Alternative approaches:
 - cladding built with deep buried oxide (thermal issues)
 - extra waveguide layers on top (process issues)

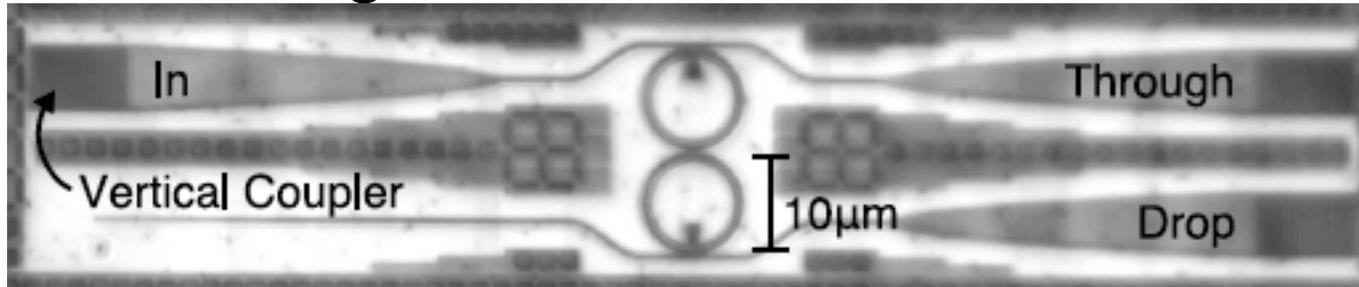
MIT Photonic Test Chip



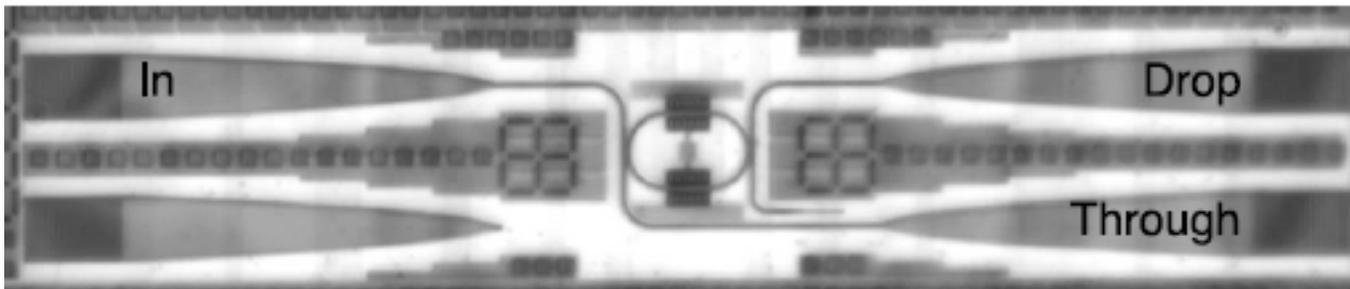
- 65nm bulk CMOS
- 114 devices
- 21 cm of waveguide

Filters/Modulators

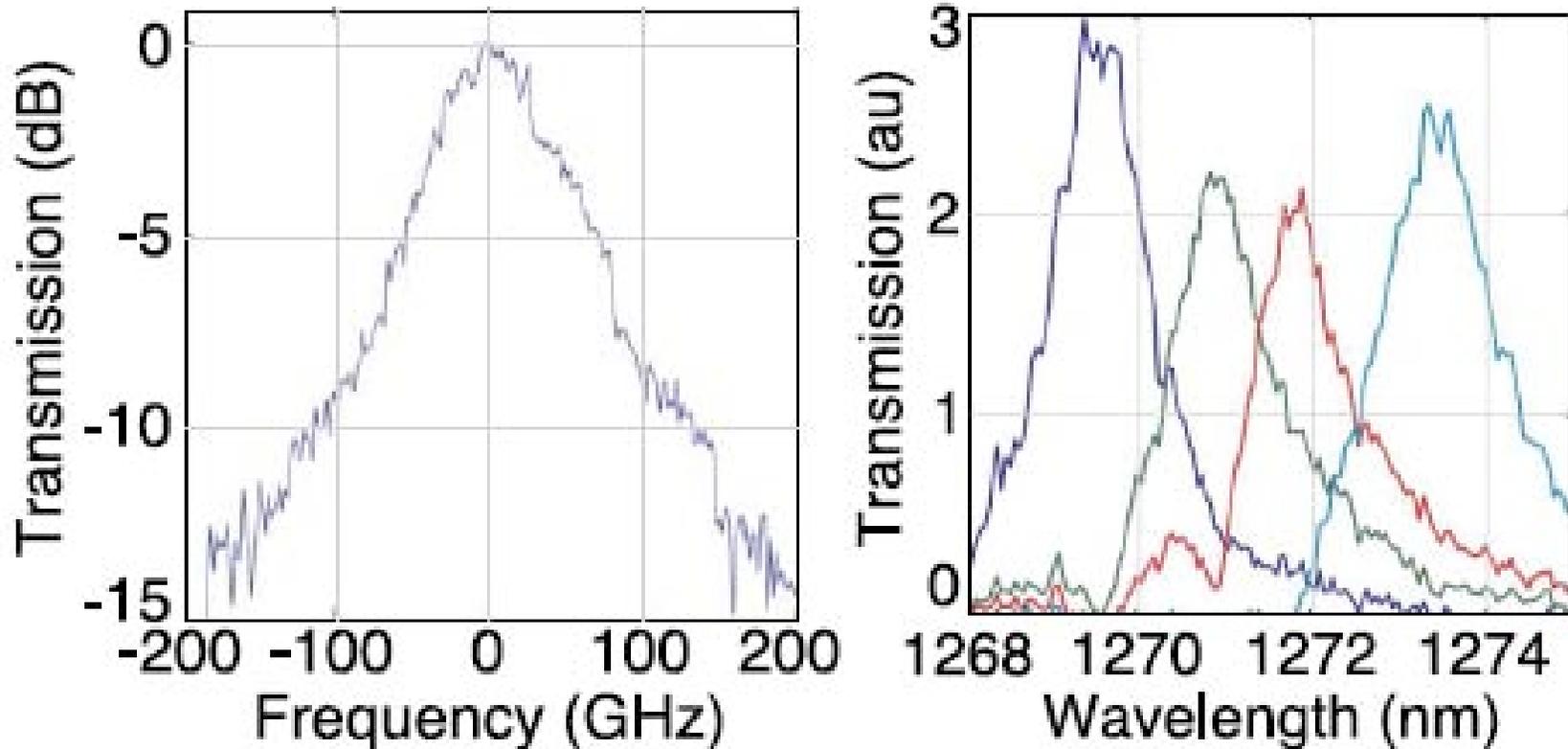
- ❑ Double-ring resonant filter



- ❑ Resonant racetrack modulator



Measured filter results



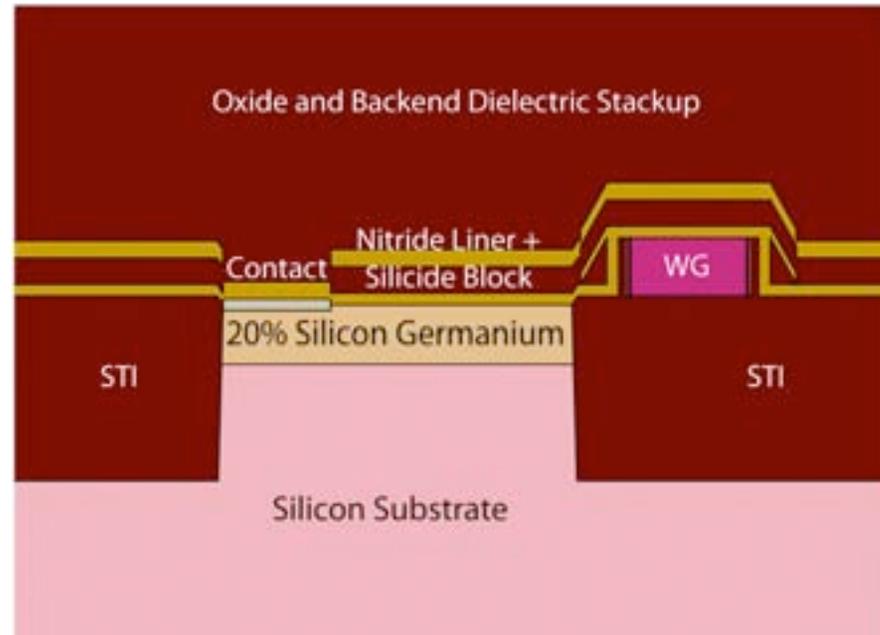
(a) Single-Ring Filter

(b) Four-way Filter Bank

Figure 3: Experimental results for single-ring filters implemented in a bulk 65 nm CMOS test chip [13]

Extrapolating from these first results, expect to improve to 64 wavelengths in each direction

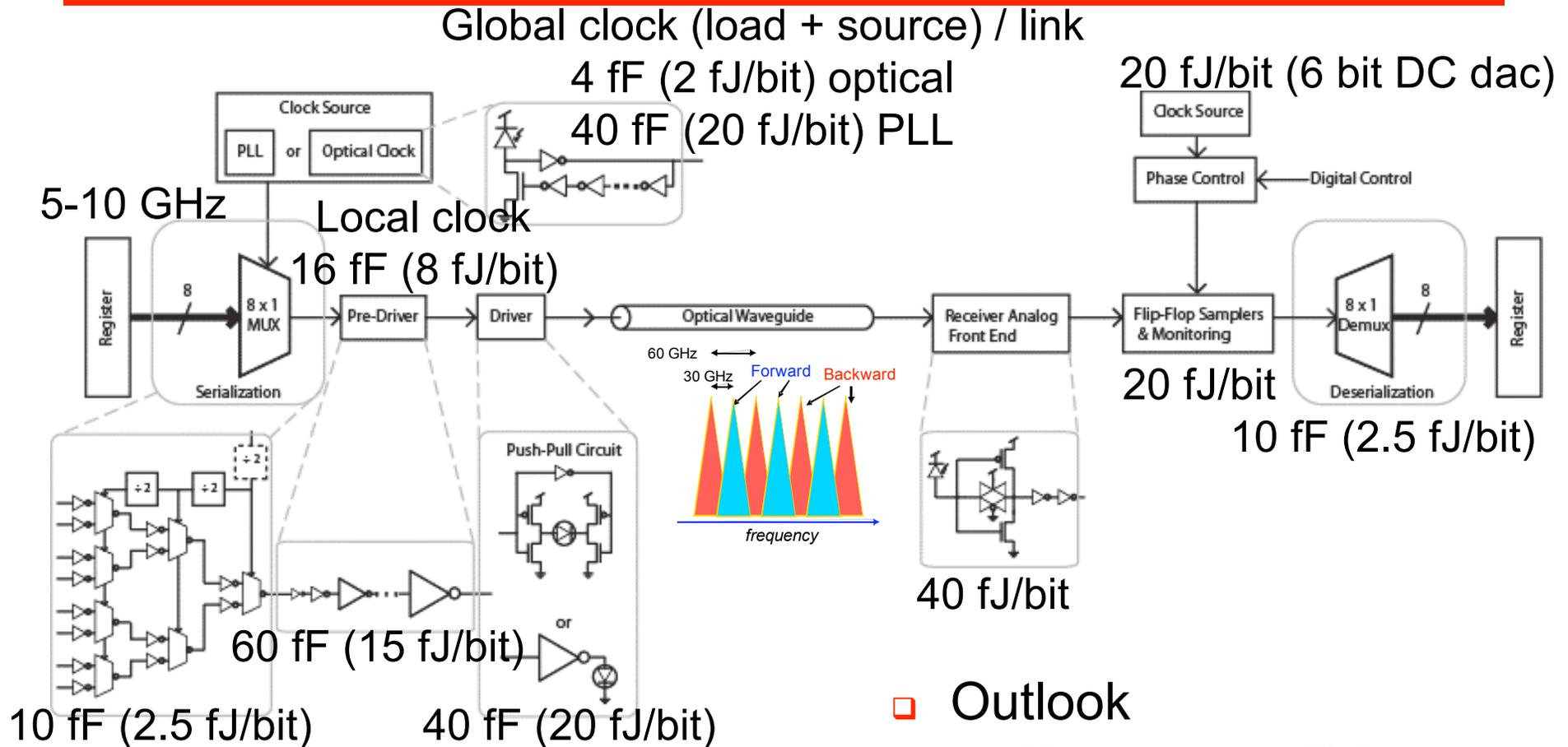
Photodetectors



Photodiode cross-section

- ❑ Embedded SiGe used to create photodetectors
- ❑ Simulation results show good capacitance (<1 fF/ μm) and dark current (<10 fA/ μm)
- ❑ Sub-100 fJ/bit energy cost required for the receiver

Photonic Link Energy Budget



Challenges

- Device/process technology
- Circuit Design
- Network Design

Outlook

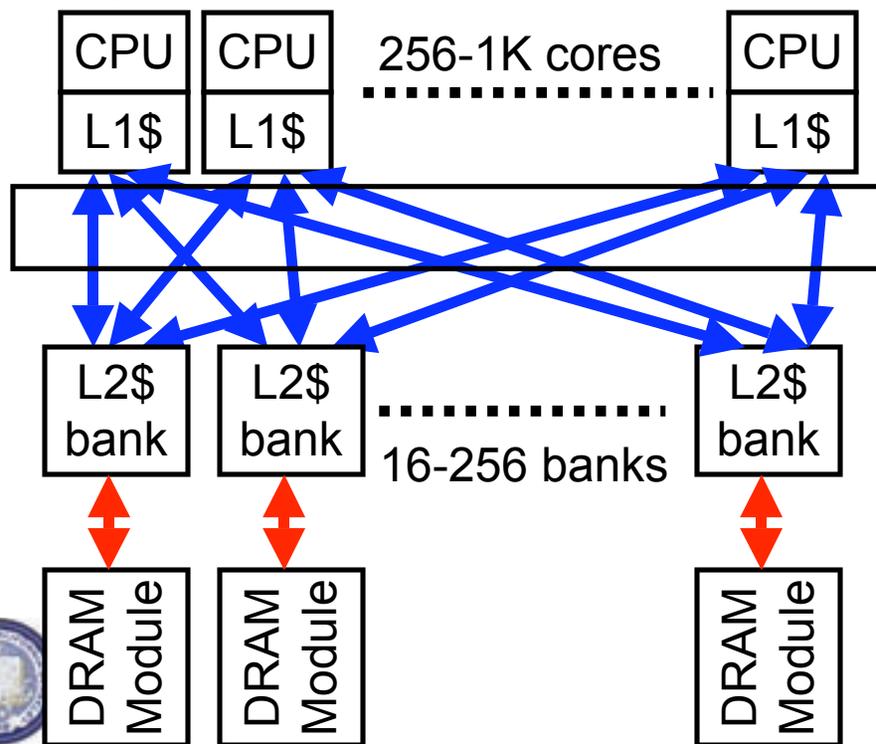
- More energy-efficient
 - < 250 fJ/bit
- Better data-rate density
 - 100+ wavelengths (1Tb/s) per waveguide



Manycore interconnect bottlenecks

- ❑ Manycore systems improve performance and power
 - Multiple simpler cores exploit thread-level and data-level parallelism
- ❑ Communication through single multi-banked shared memory
 - Simplifies programming and improves performance

Manycore Microprocessor



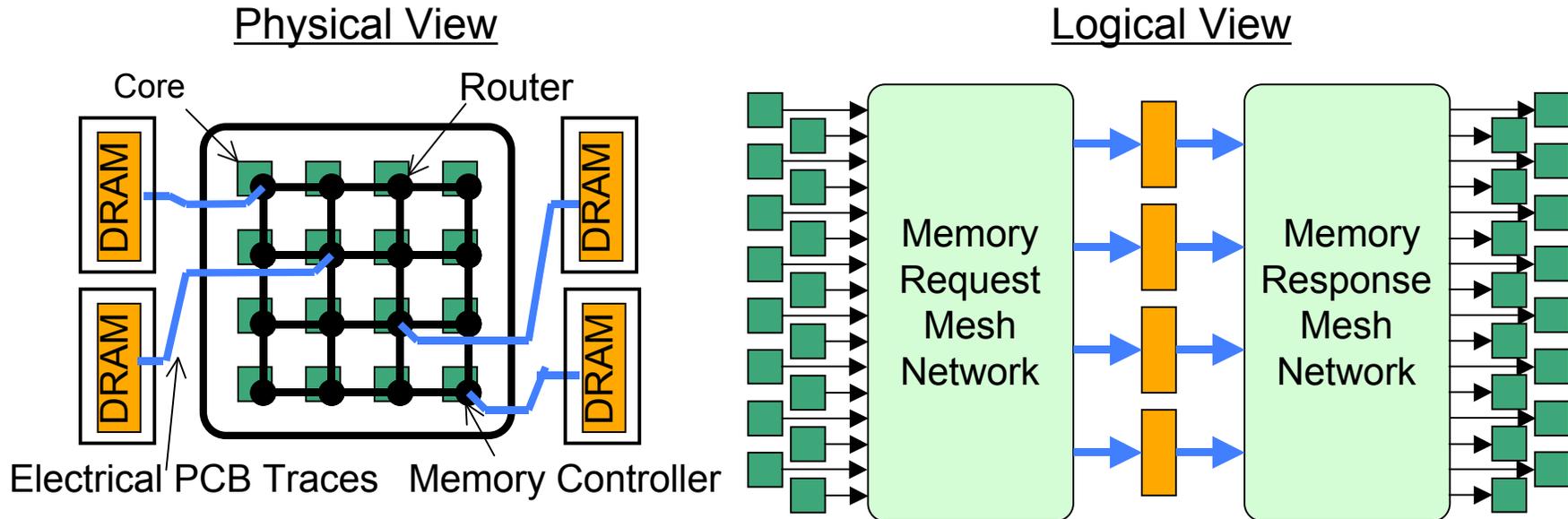
- ❑ Shared memory system power and performance bottlenecks

- Core-core connectivity via shared outer-level cache banks
- Outer cache to off-chip DRAM



Electrical system architecture

(L2 cache ignored for now, direct core-to-DRAM network shown)

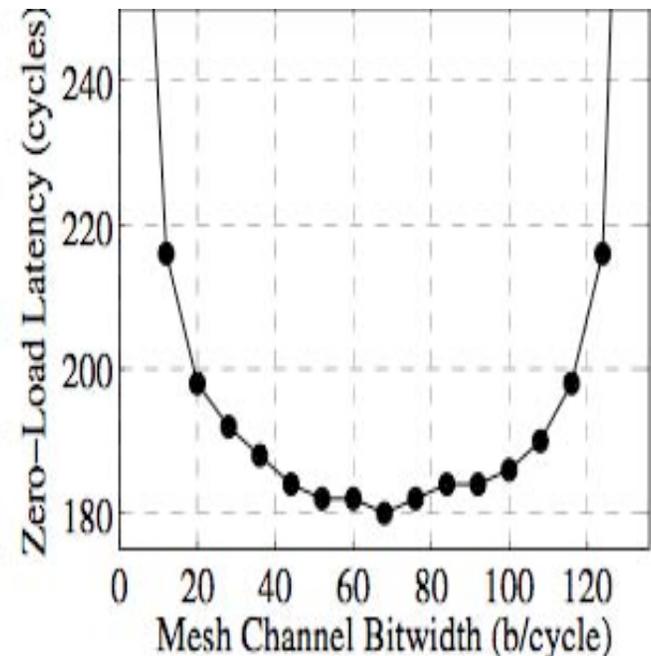
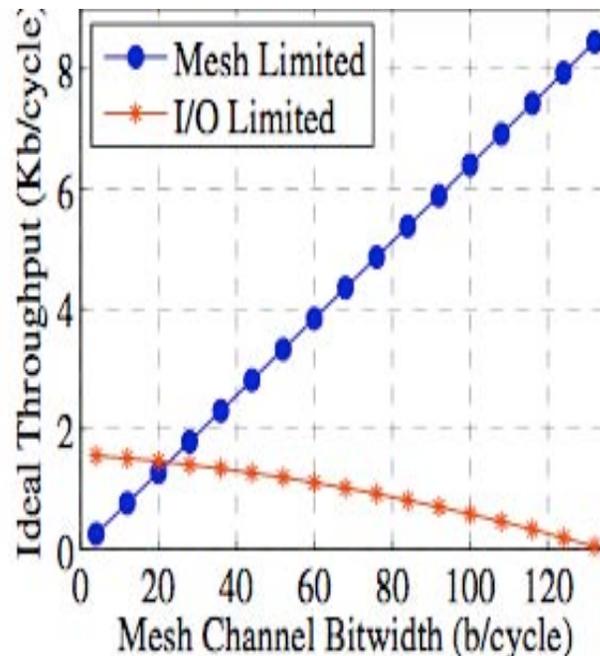
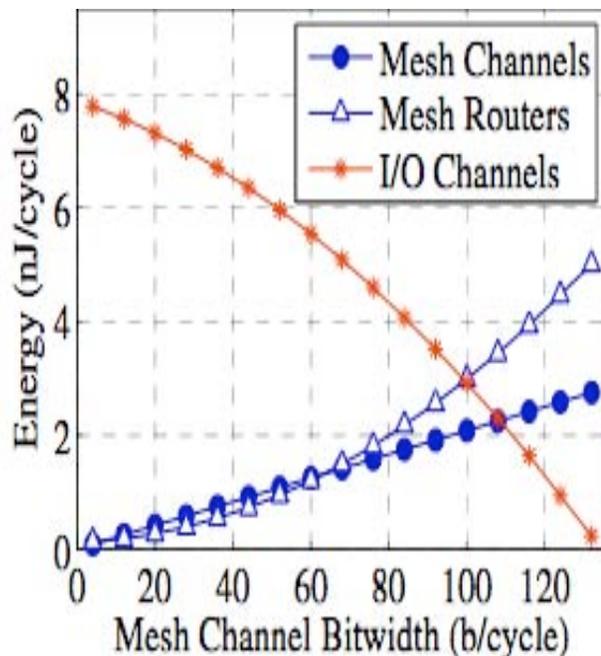


- ❑ Two on-chip electrical mesh networks
 - Request network for sending requests to the appropriate memory controller
 - Response network for sending responses back to the appropriate core
- ❑ Memory controllers distributed across chip
 - Flip-chip I/O and standard PCB traces to communicate with DRAM modules
 - Package pin limit and electrical signaling limits memory bandwidth
 - Limited chip power budget also limits memory bandwidth

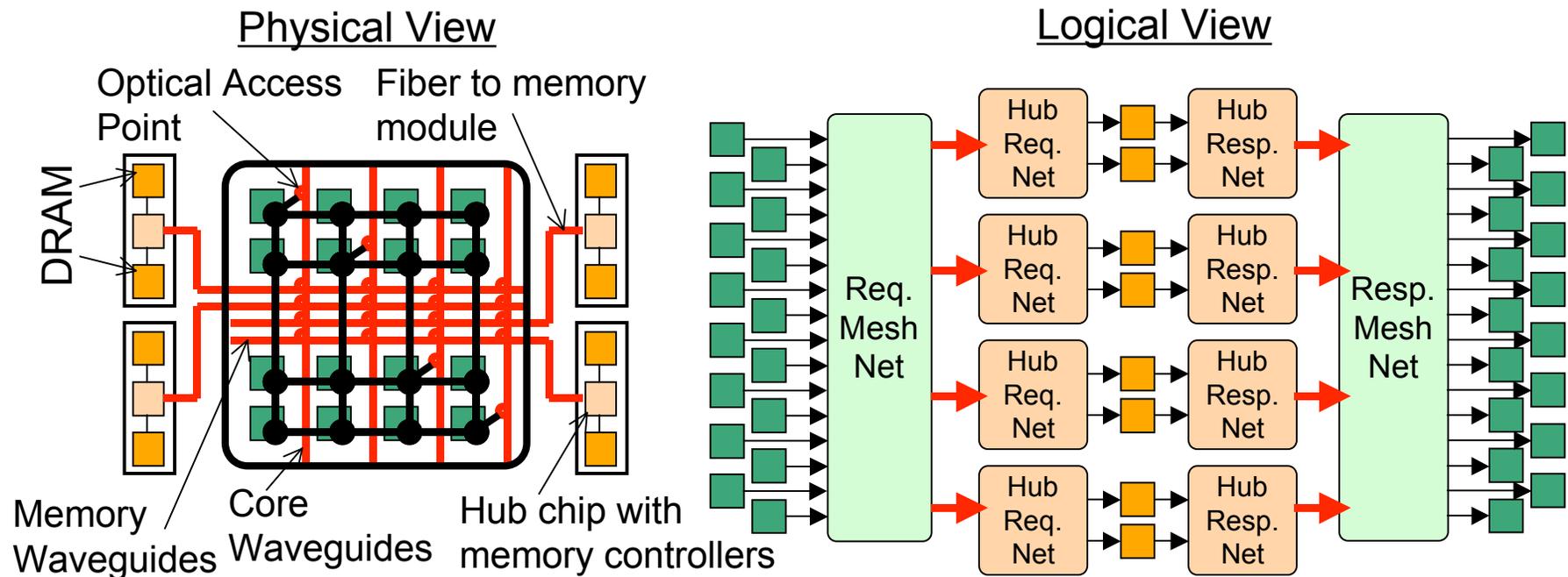


Power-constrained design exploration

- Analytical model of on-chip wires, routers, and off-chip I/O
 - 256 cores at 2.5GHz in predicted 22nm technology
 - Energy constrained to 8nJ/cycle (20W)
 - Uniform random traffic pattern
- Vary bitwidth of channels between on-chip mesh routers
 - Wider channels improve on-chip bandwidth but require more energy thereby reducing off-chip I/O bandwidth



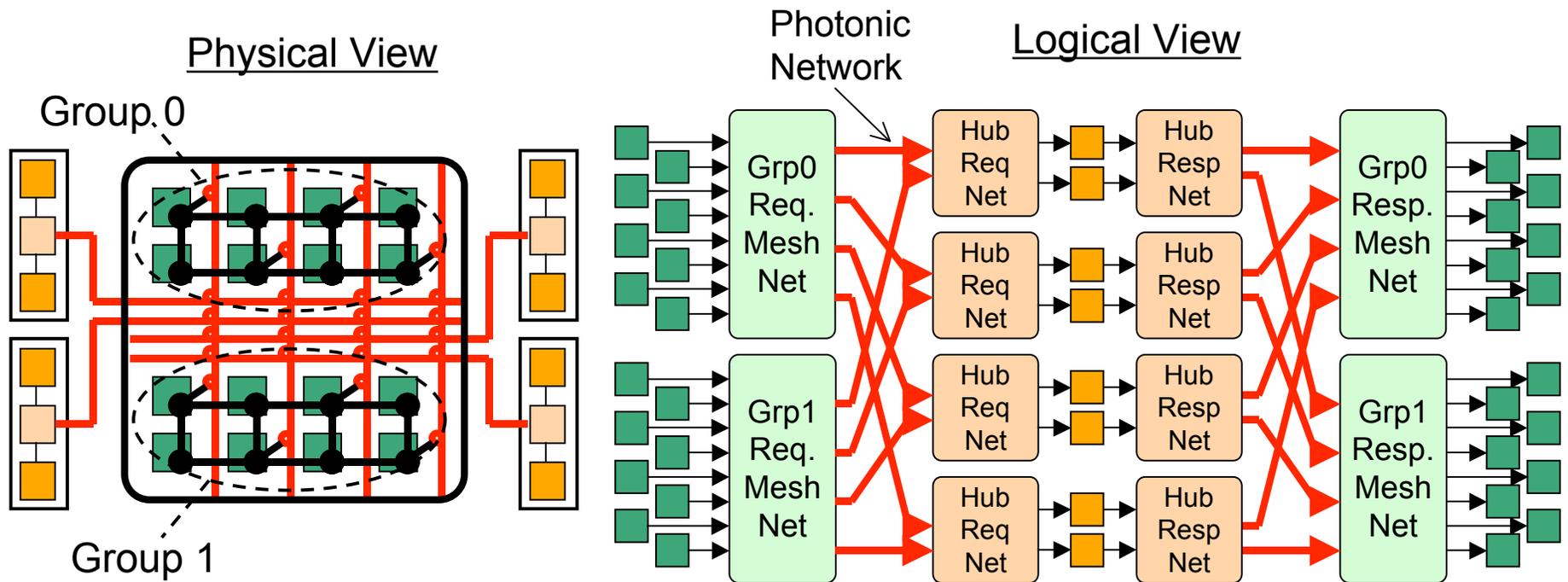
Photonic system architecture



- ❑ Memory request message path
 - Electrical mesh to reach appropriate photonic access point
 - Core waveguide to photonic switch matrix
 - Statically routed through photonic switch matrix
 - Memory waveguide and optic fiber to reach hub chip
 - Routed through hub chip to correct DRAM chip
 - (Separate photonic/mesh networks carry responses back to core)
- ❑ *Removes I/O pin/energy bottleneck, but on-chip electrical mesh now becomes energy bottleneck*



Photonic multi-group system architecture

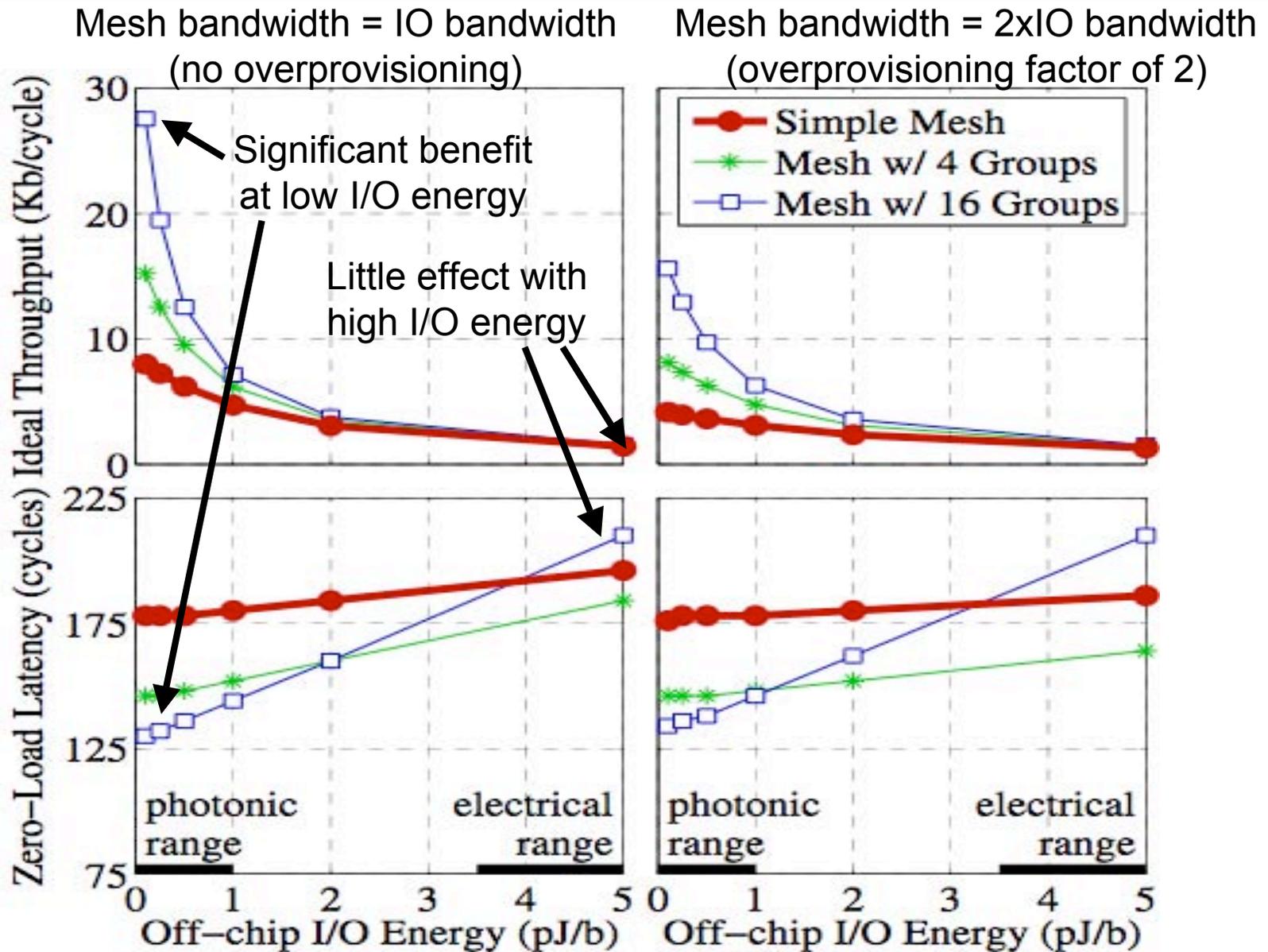


- Divide cores into groups, each with local mesh (no global mesh)
- Each group still has one AP for each DIMM and thus can access all of memory
- Since there are more APs, each AP is narrower (uses less λ s)
- Hub chip now include a crossbar network and arbitration for DRAMs
- Uses photonic network as a very efficient global crossbar

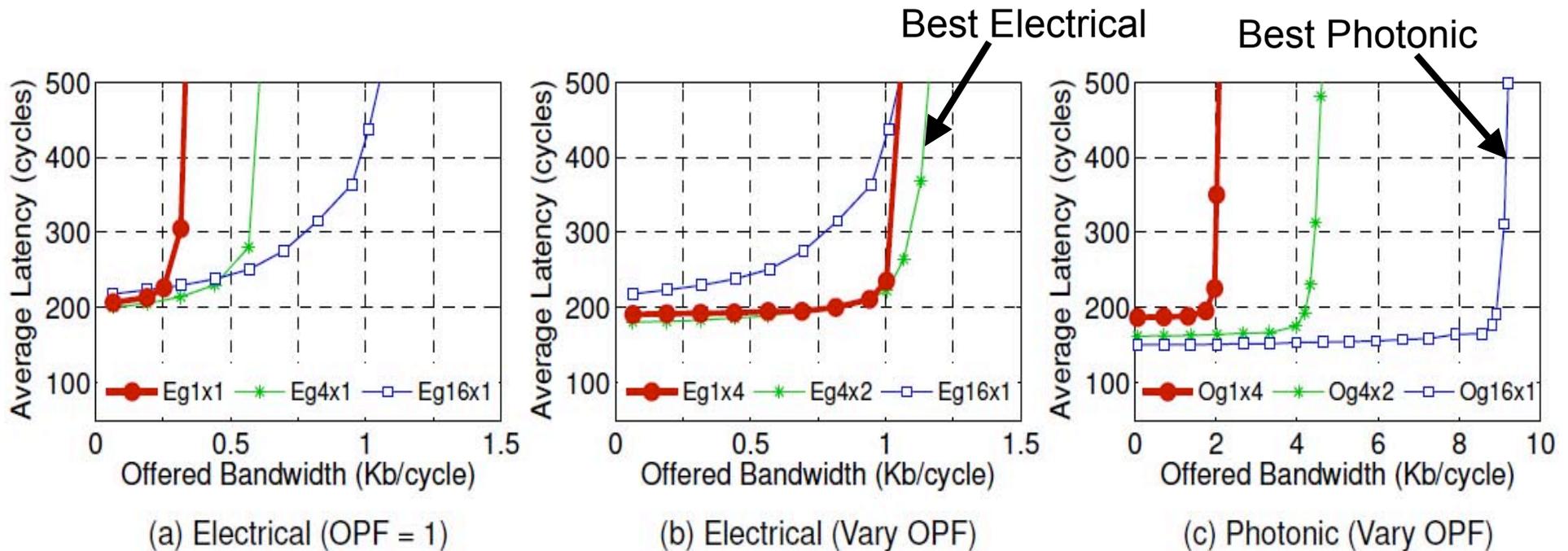


Seamless cross-chip/off-chip photonic network alleviates interconnect energy bottleneck

Analytical Analysis of Groups



Detailed Simulation Evaluation



- ❑ Model includes pipeline latencies, router contention, message fragmentation, credit-based flow control, and serialization overheads
- ❑ Uniform random addresses

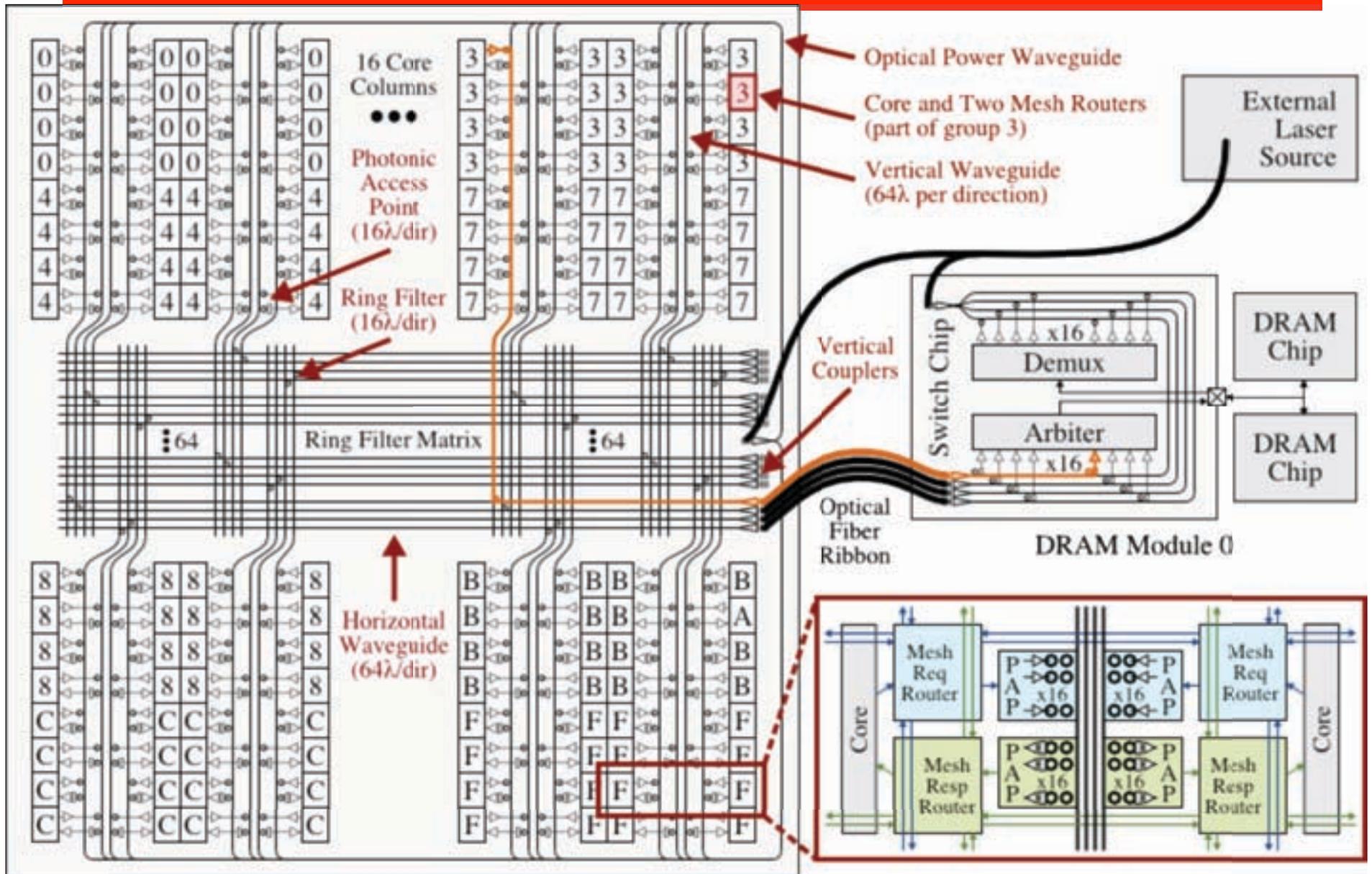
Simulation Parameters

- ❑ 256 cores, 16 DRAM modules
- ❑ Mesh routers 2 cycles
- ❑ Mesh channels 1 cycle
- ❑ Global crossbar 8 cycles
- ❑ DRAM access 100 cycles (40ns)

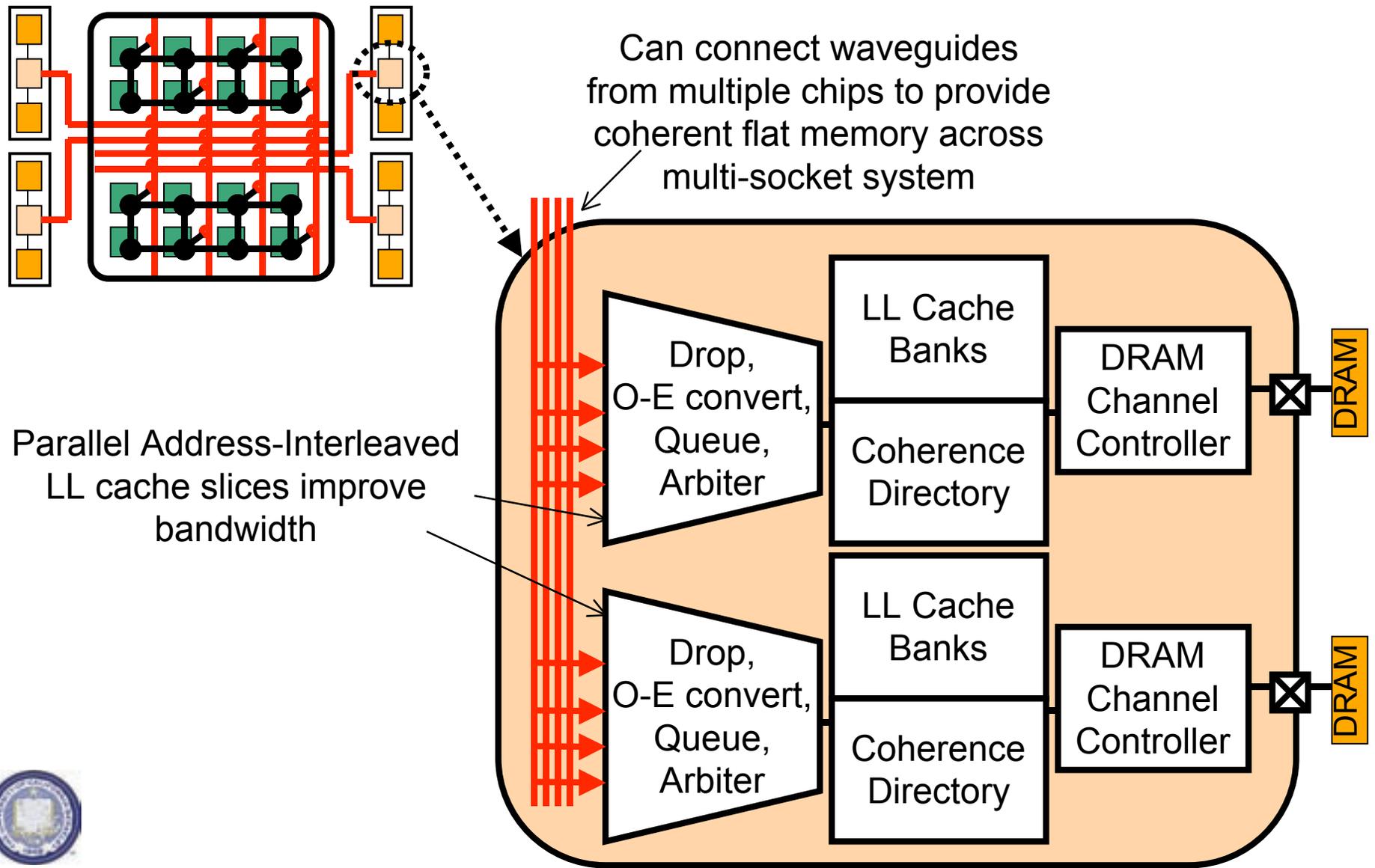


Best photonic is approximately 9x better bandwidth than best electrical at same energy constraint, with lower latency

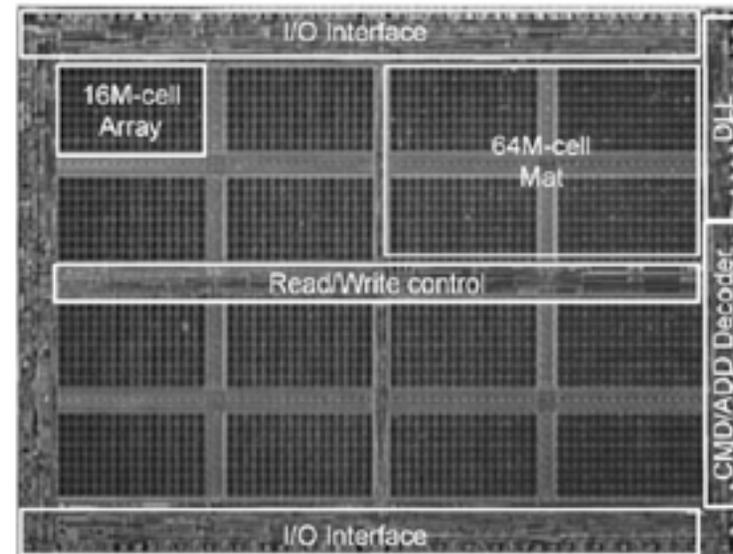
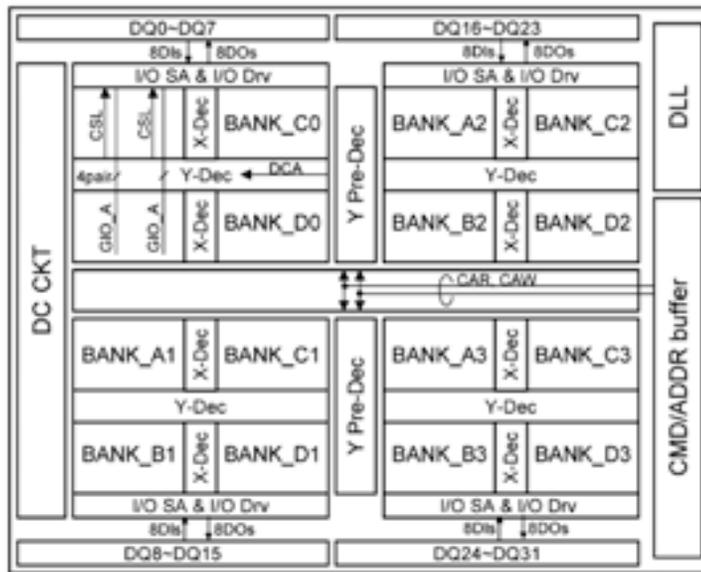
Full System (Conventional DRAM)



Last-Level Cache at DRAM switch



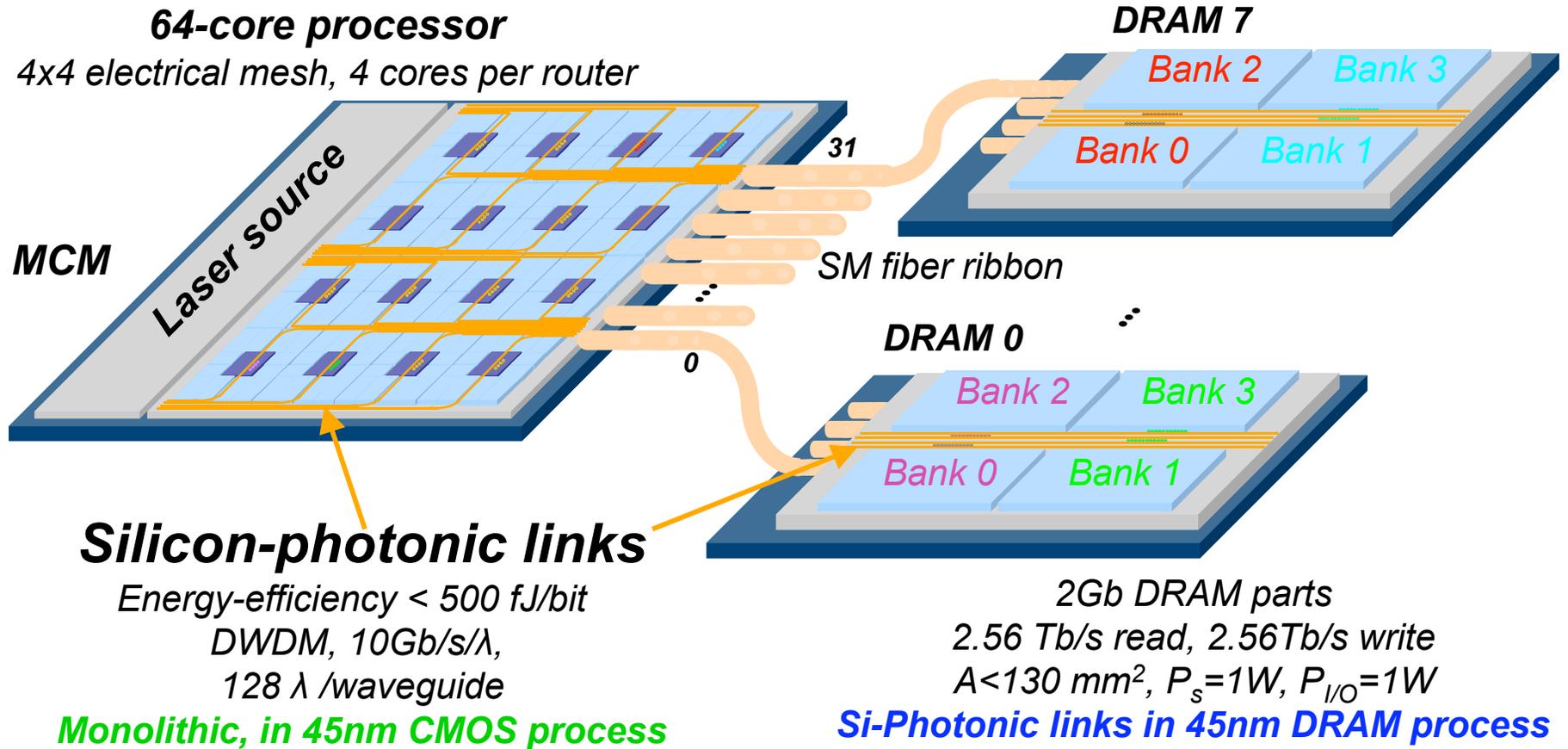
Improving DRAM organization for high throughput



- ❑ Photonics-to-DRAM offers very high throughputs
 - 1-2 Tb/s/fiber
 - 5-10 Tb/s fits within DRAM I/O power budget
 - But, need to organize banks and switching to support this high throughput



Optically-banked DRAM

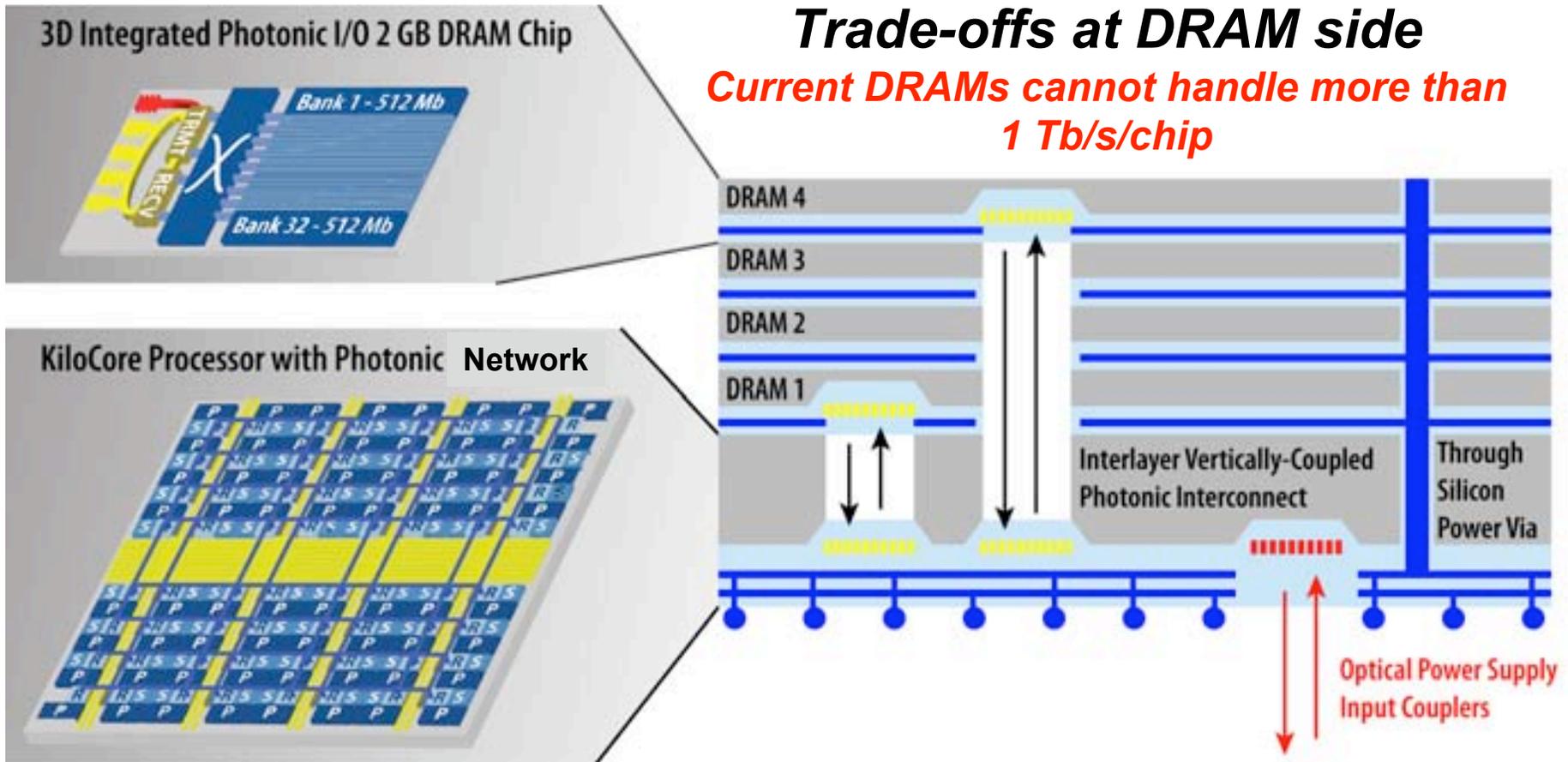


Si-Photonic Interconnect in Embedded Aps

Electrical vs. Optical interconnect

Trade-offs at DRAM side

Current DRAMs cannot handle more than 1 Tb/s/chip



Electrical vs. Optical interconnect

Trade-offs at Processor side

Current Processors have less than 2 Tb/s throughput

Interesting packaging options for optics with vertical coupling and TSVs



Acknowledgements

□ Students

- Christopher Batten
- Michael Georgas
- Charles Holzwarth
- Ajay Joshi
- Anatoly Khilo
- Daniel Killibrew
- Jin Kwon
- Jonathan Leu
- Hanqing Li
- Benjamin Moss
- Jason Orcutt
- Miloš Popović
- Imran Shamim
- Jie Sun

[For more details see our Hot Interconnects 2008 paper]

□ Support

- Texas Instruments
- DARPA MTO

