Design-Space Exploration for CMOS Photonic Processor Networks

Vladimir Stojanović[1], Ajay Joshi[2], Cristopher Batten[3], Yong-Jin Kwon[4], Scott Beamer[4], Sun Chen[1] and Krste Asanović[4]

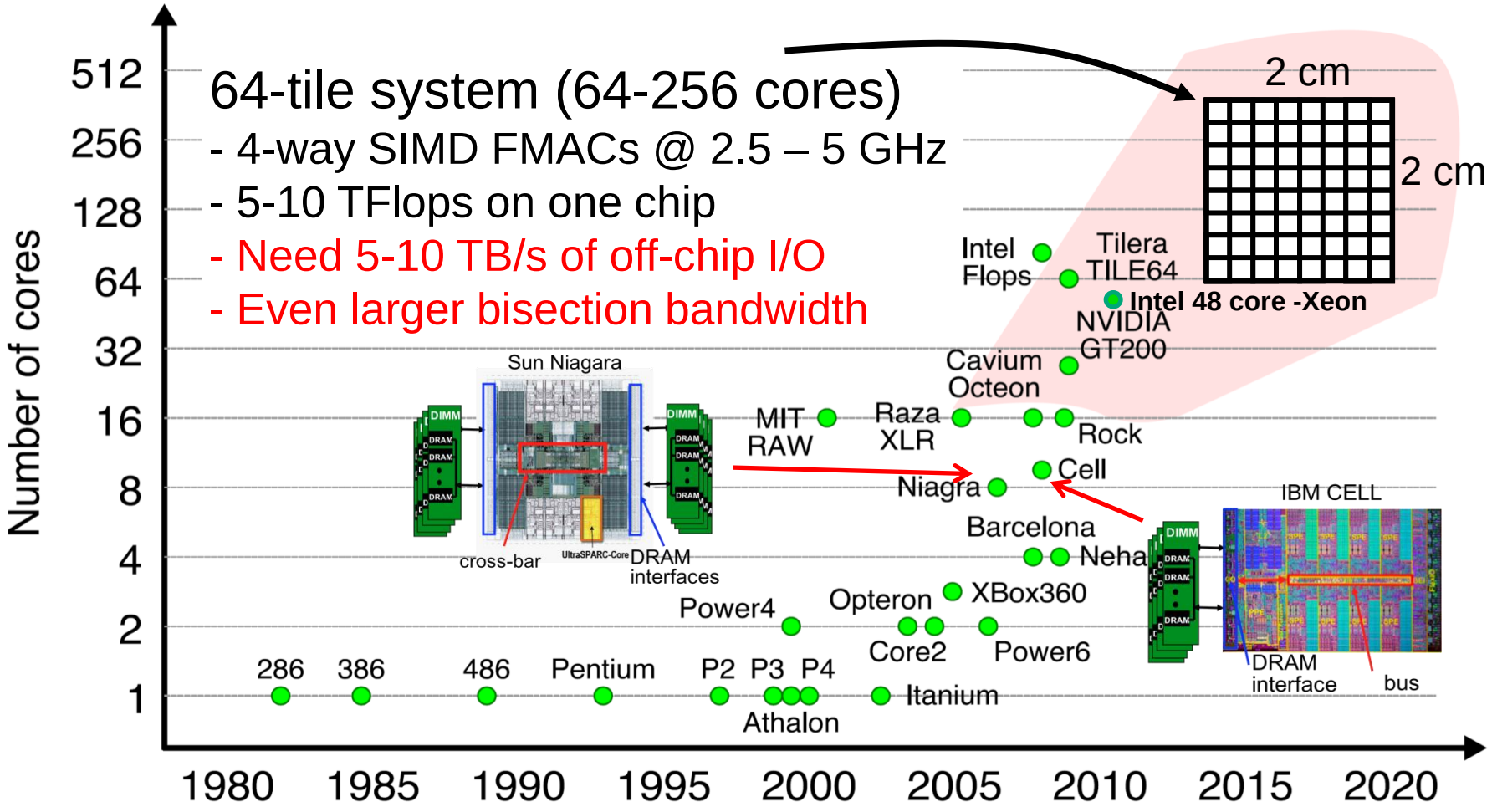
[1]MIT, [2]Boston University,
[3]Cornell University, [4]UC Berkeley

Acknowledgments

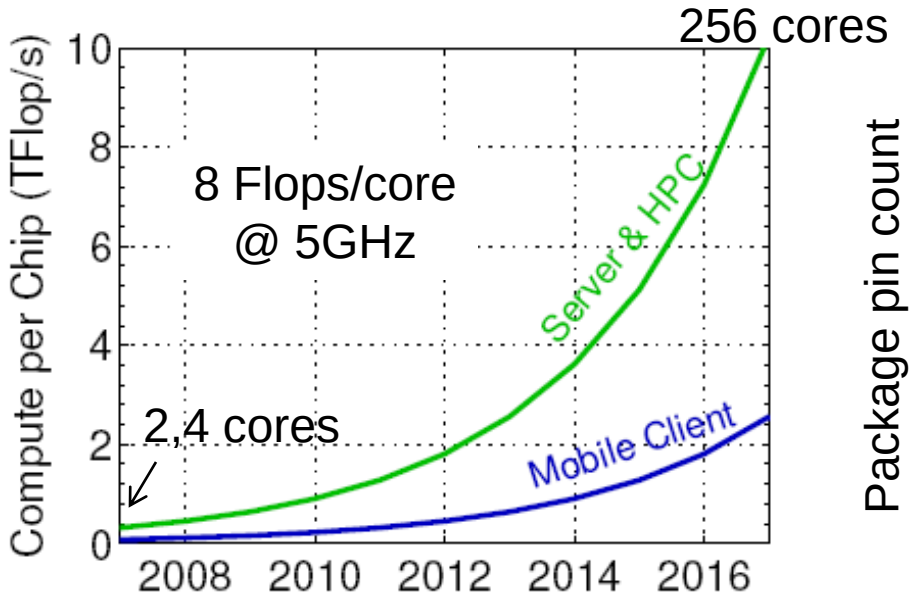
- Rajeev Ram, Milos Popovic, Franz Kaertner, Judy Hoyt, Henry Smith, Erich Ippen
- Hanqin Li, Charles Holzwarth
- Jason Orcutt, Anatoly Khilo, Ben Moss, Jie Sun, Jonathan Leu, Michael Georgas, Imran Shamim

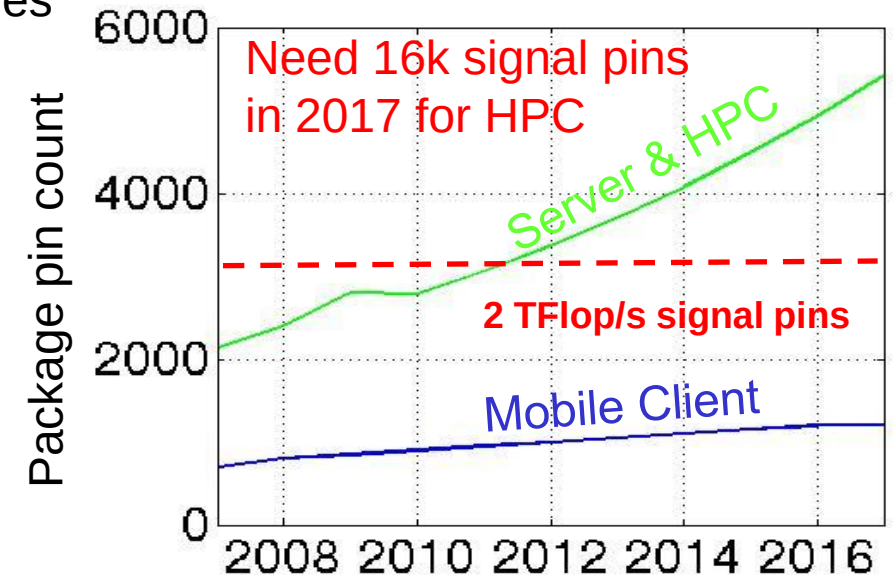
- Dr. Jag Shah – DARPA MTO
- Texas Instruments
- Intel Corporation

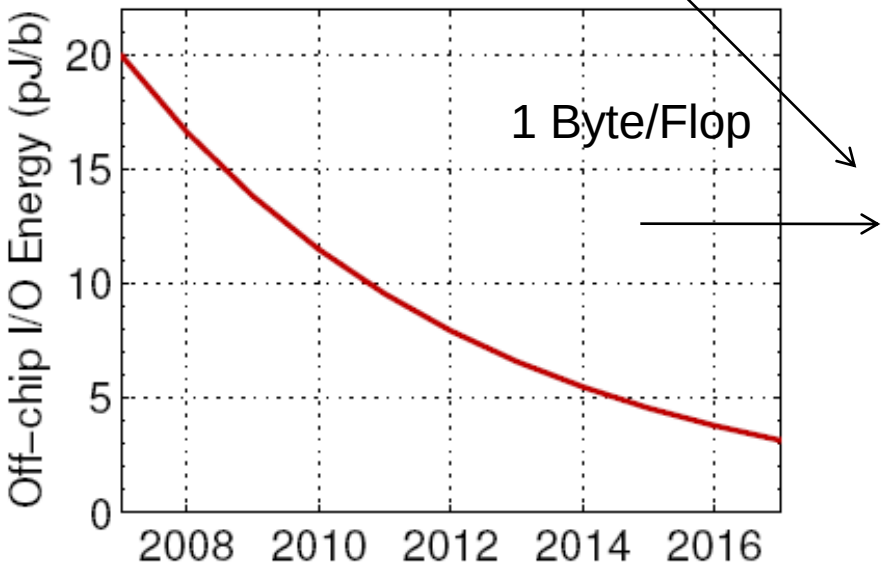
Processors scaling to manycore systems

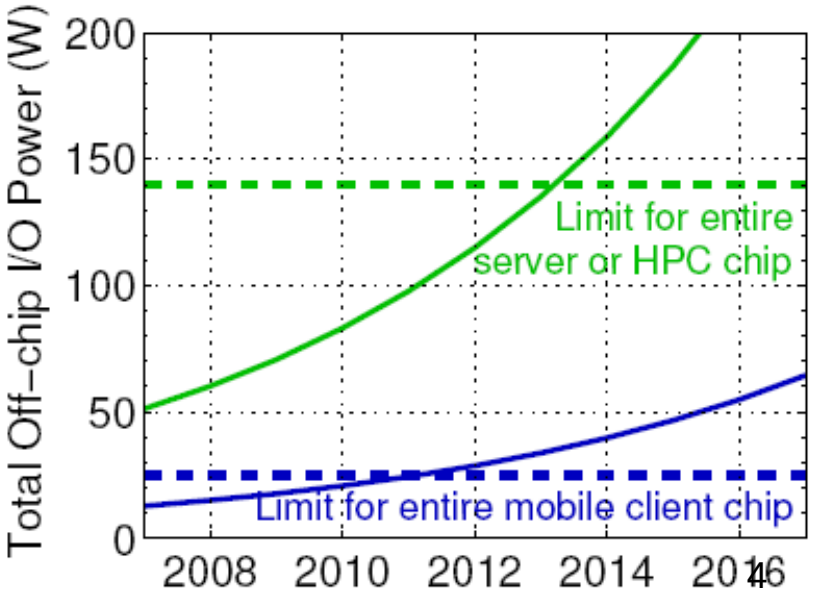


Bandwidth, pin count and power scaling

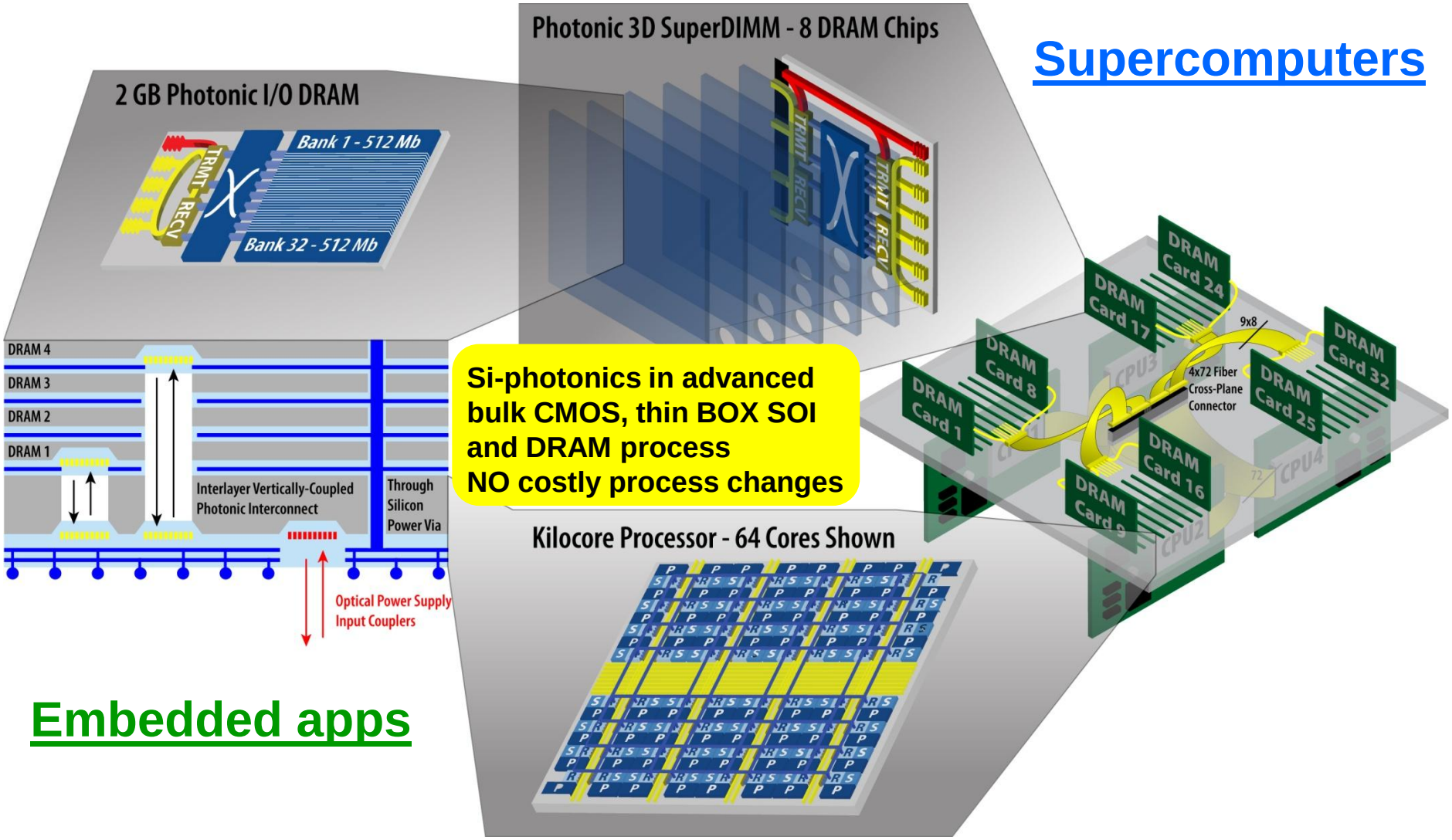








Monolithic CMOS-Photonics in Computer Systems



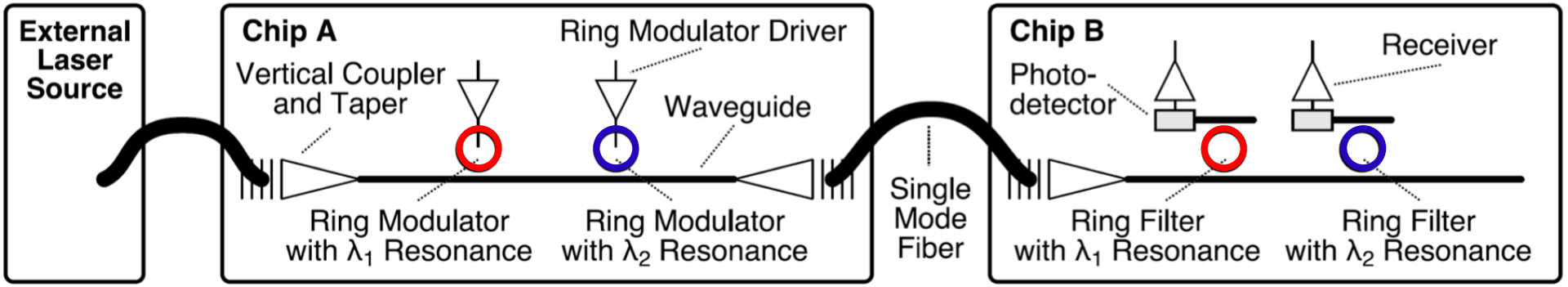
Supercomputers

Embedded apps

Bandwidth density – need dense WDM

Energy-efficiency – need monolithic integration

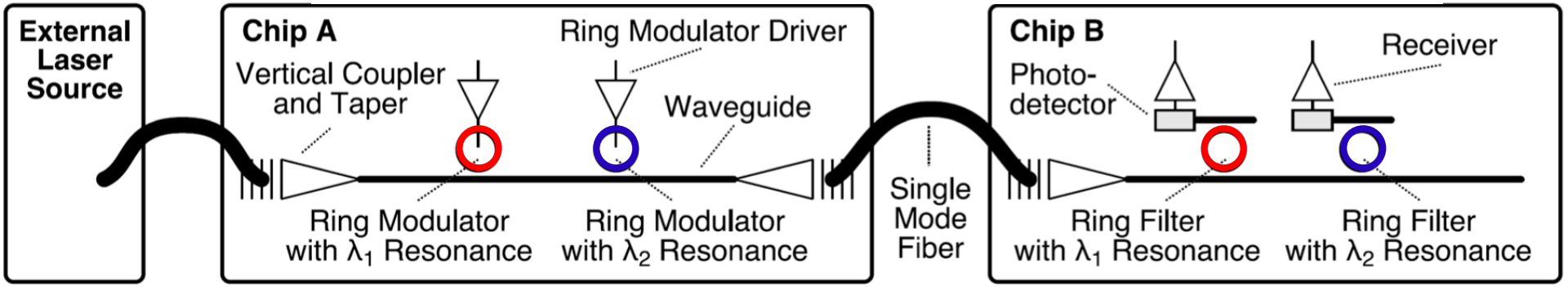
CMOS photonics density and energy advantage

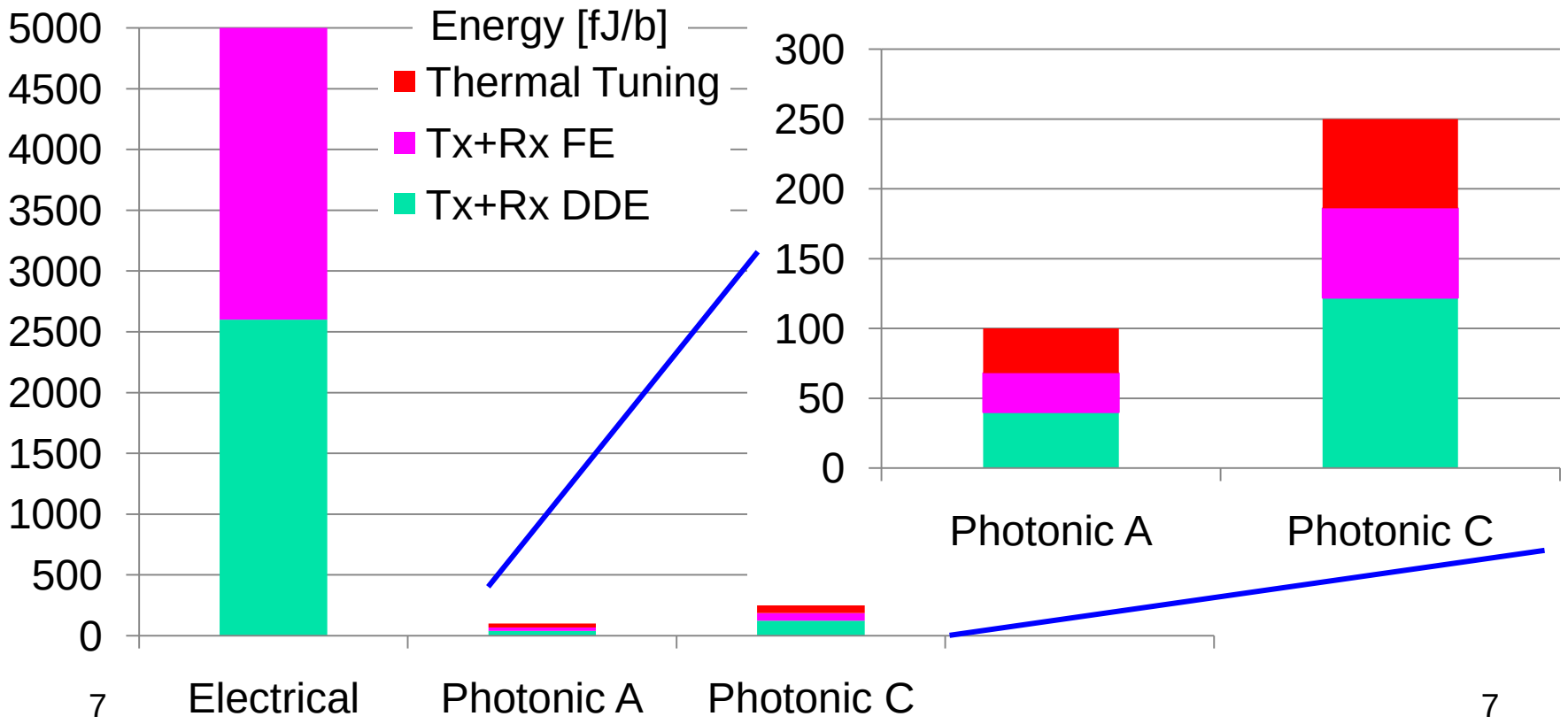


Metric	Energy (pJ/b)	Bandwidth density (Gb/s/μ)
Global on-chip photonic link	0.25	160-320
Global on-chip optimally repeated electrical link	1	5
Off-chip photonic link (100 μ coupler pitch)	0.25	6-13
Off-chip electrical SERDES (100 μ pitch)	5	0.1

But, need to keep links fully utilized …

Fixed and static energy increase at low link utilization !

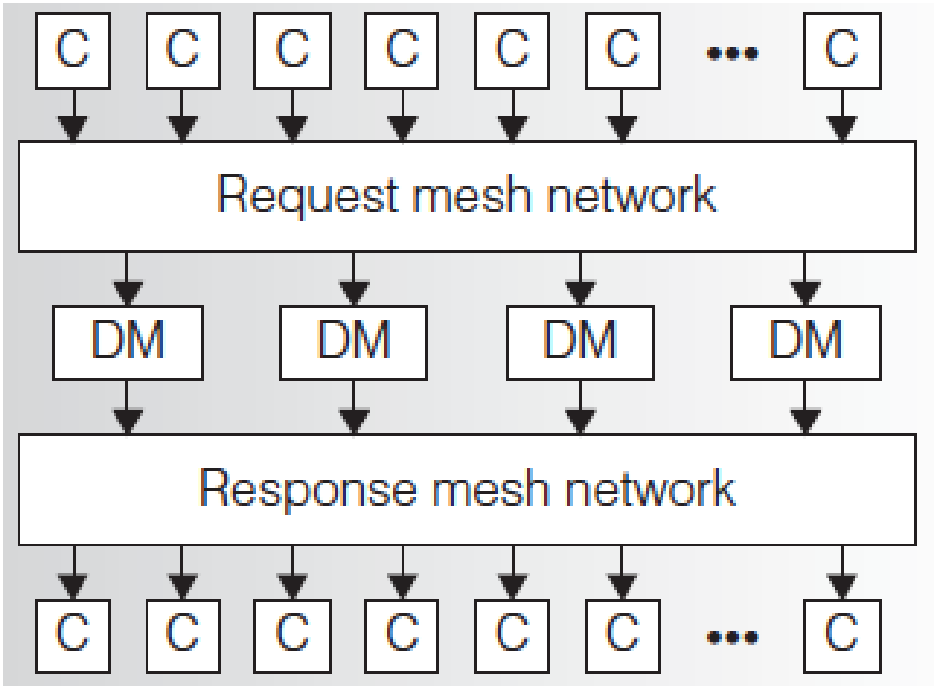


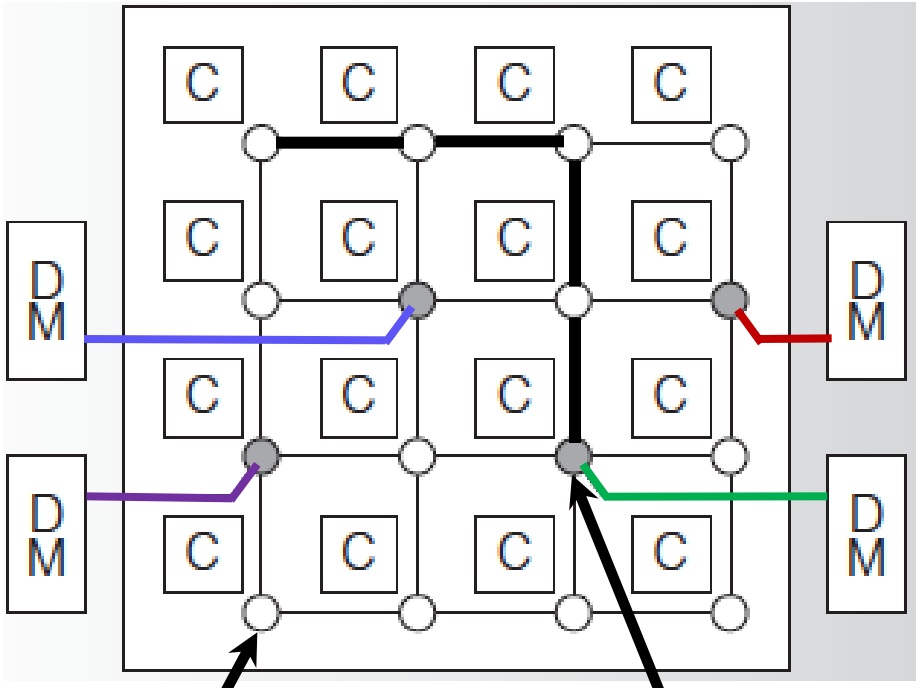


7

Core-to-Memory network: Electrical baseline

C = Core, DM = DRAM Module

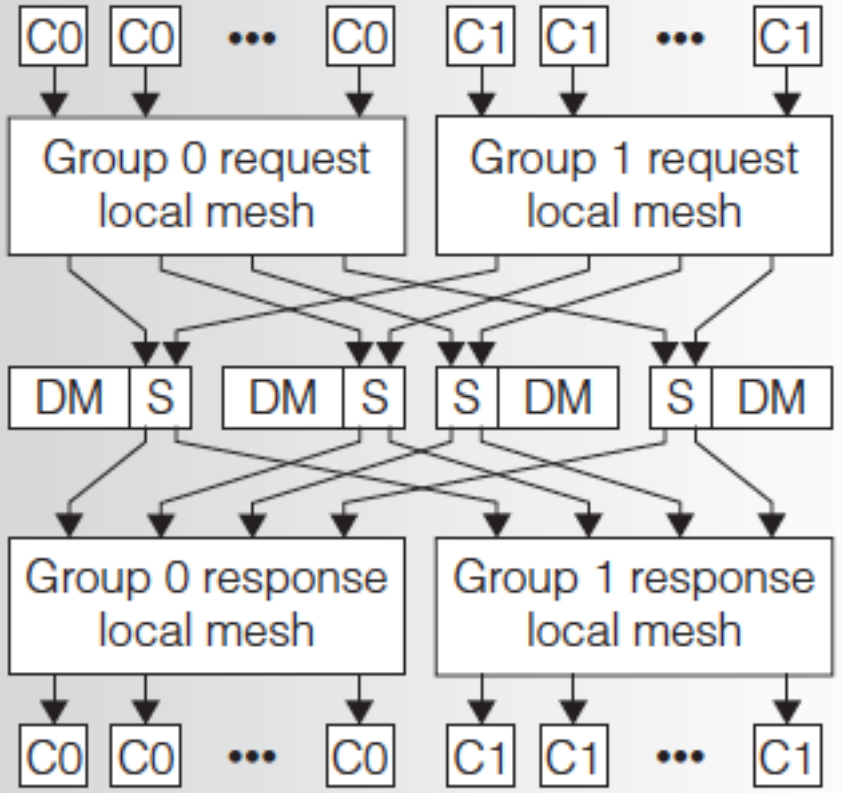


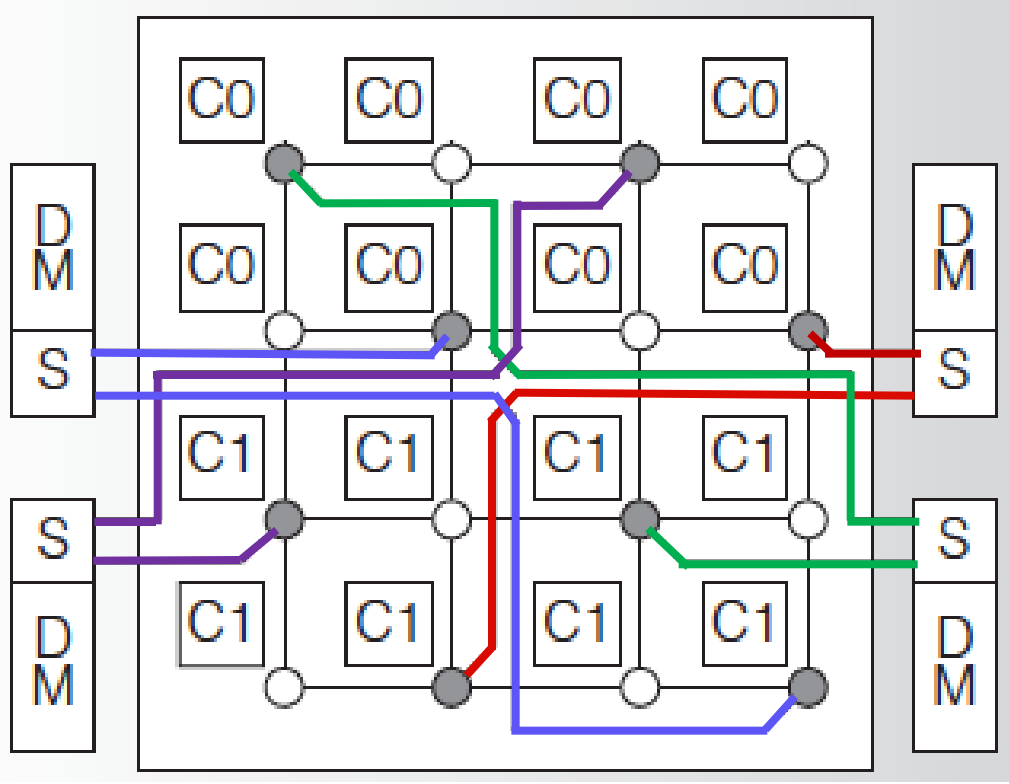


- Both cross-chip and I/O costly

Aggregation with Optical LMGS* network

* Local Meshes to Global Switches





Ci = Core in Group i, DM = DRAM Module, S = Crossbar switch

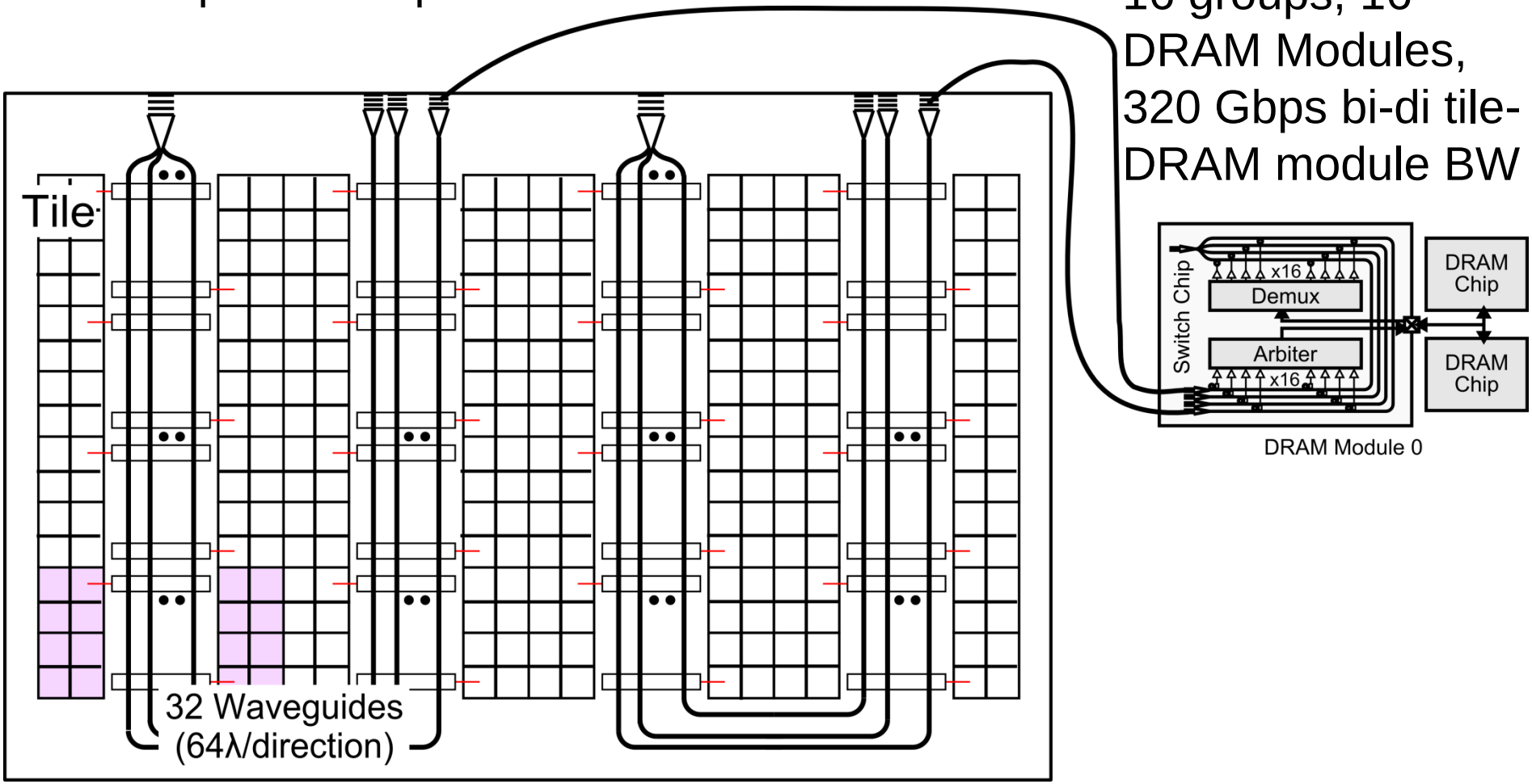
- Shorten cross-chip electrical
- Photonic both part cross-chip and off-chip

9

Photonic LMGS: Physical Mapping

Network layout optimization significantly affects the component requirements

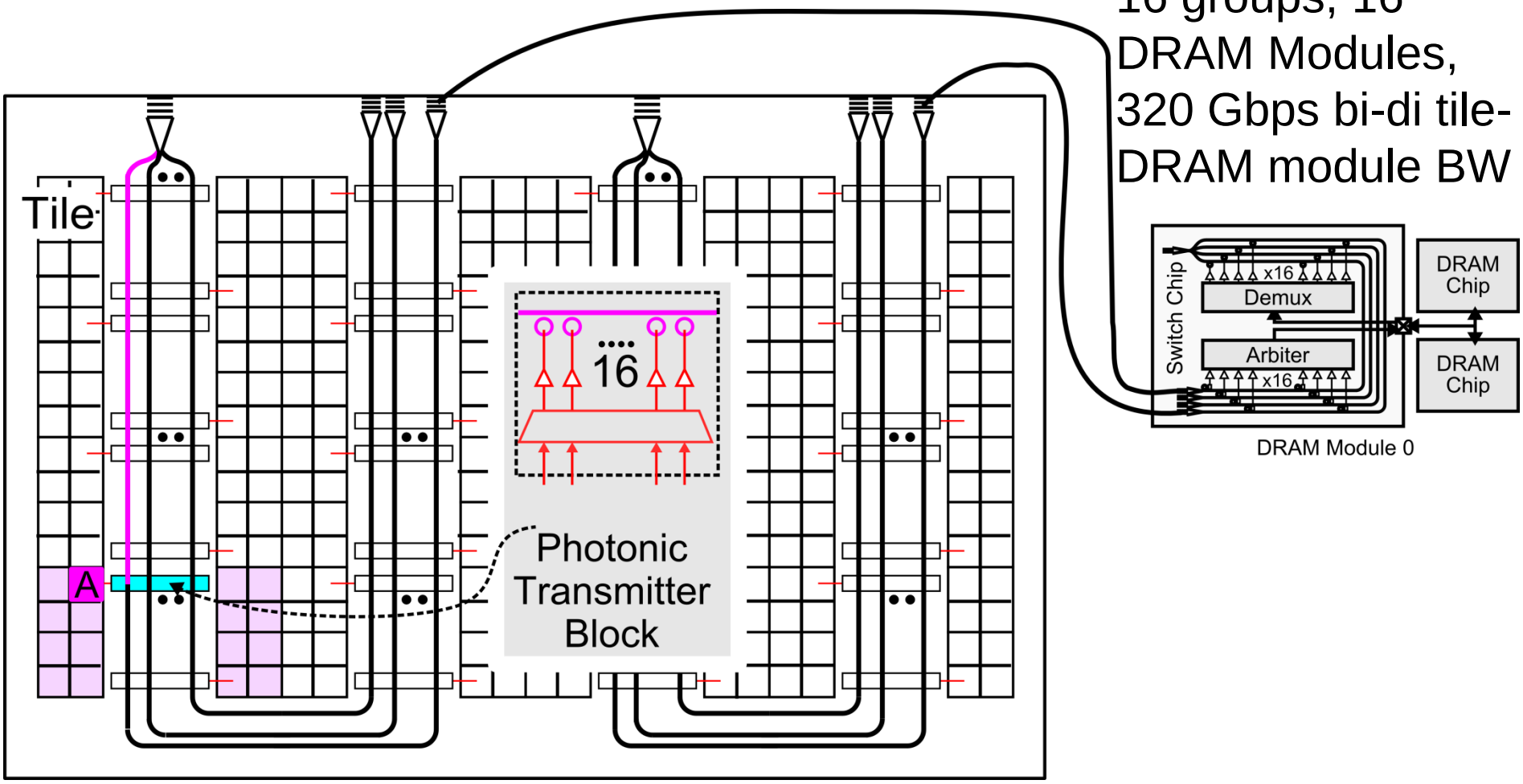
64-tile system w/ 16 groups, 16 DRAM Modules, 320 Gbps bi-di tile-DRAM module BW



[Joshi et al – PICA 2009]

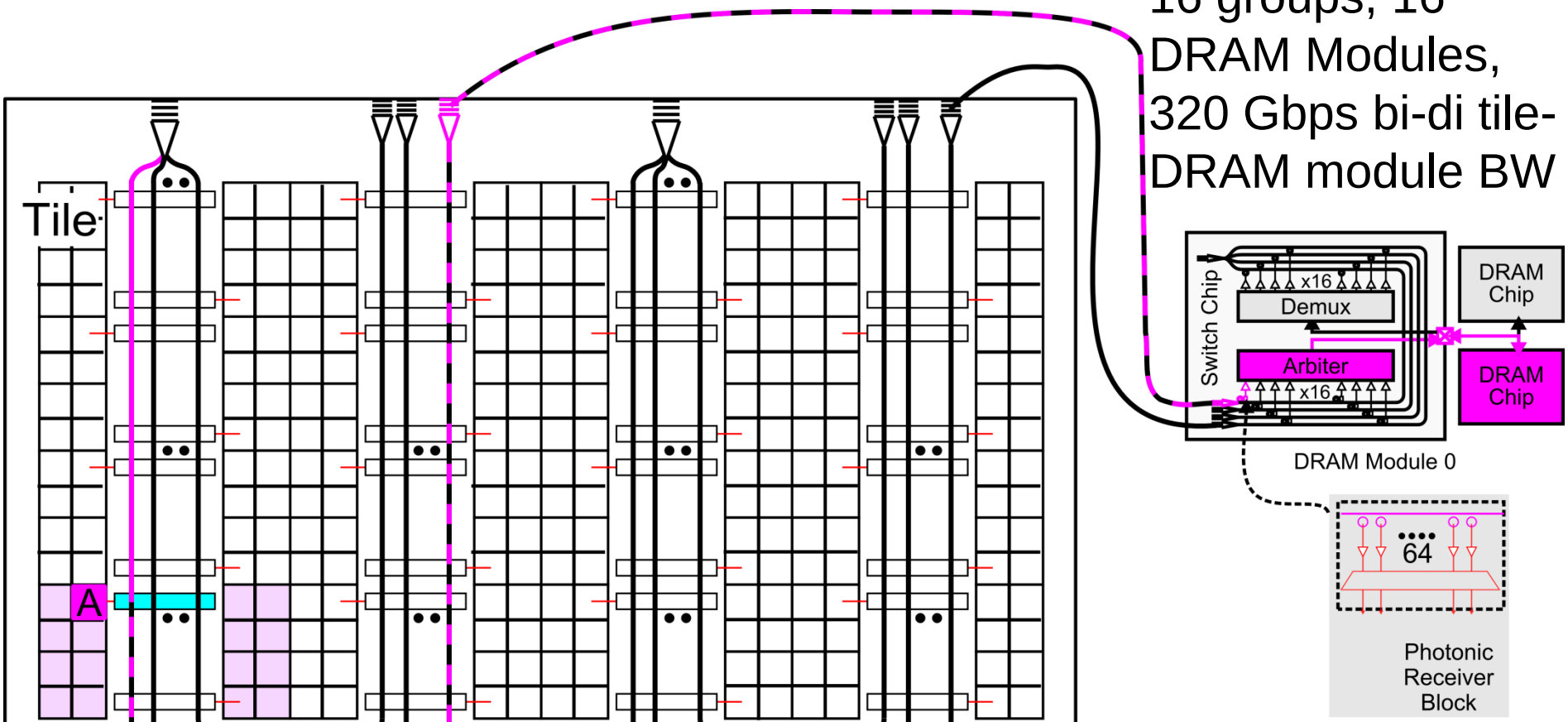
Photonic LMGS - U-shape

64-tile system w/ 16 groups, 16 DRAM Modules, 320 Gbps bi-di tile-DRAM module BW



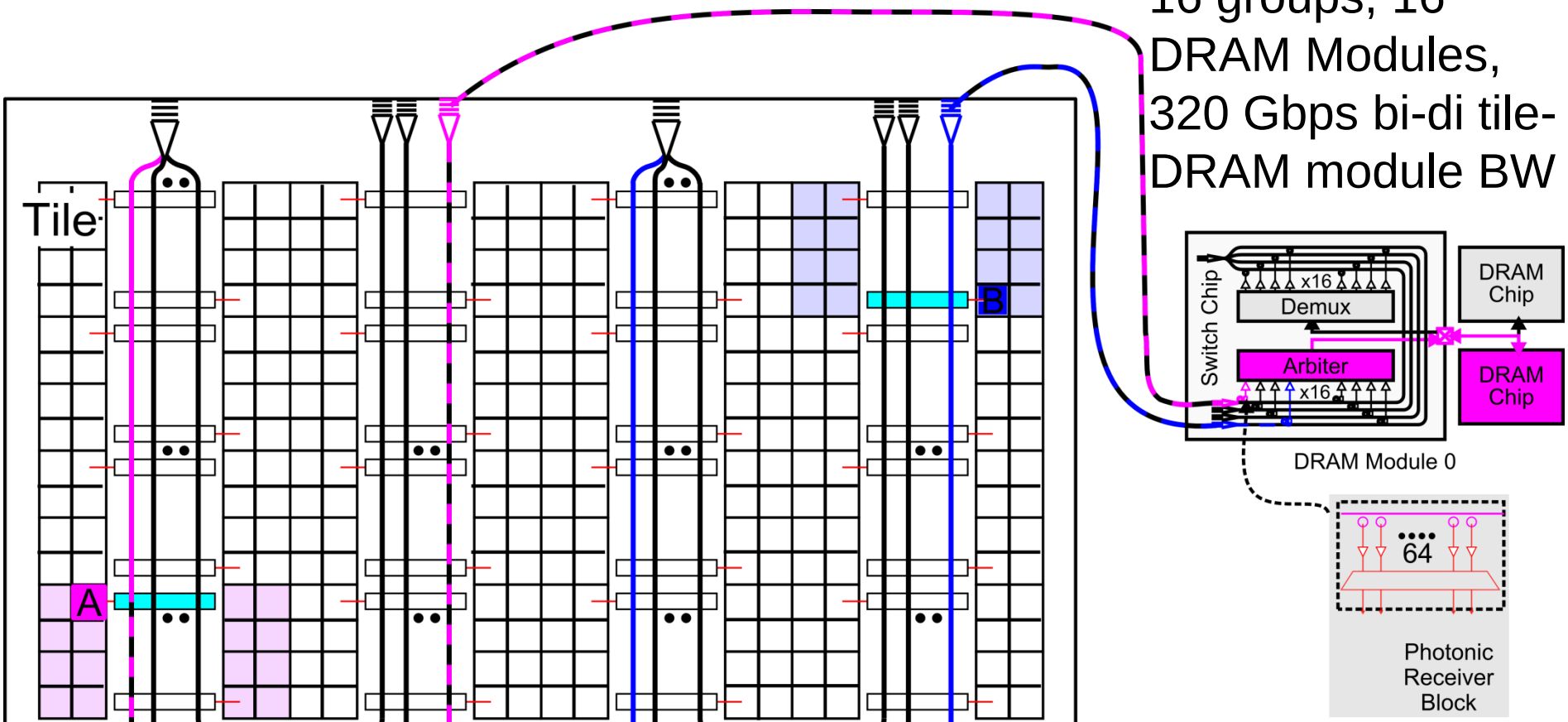
Photonic LMGS - U-shape

64-tile system w/ 16 groups, 16 DRAM Modules, 320 Gbps bi-di tile-DRAM module BW



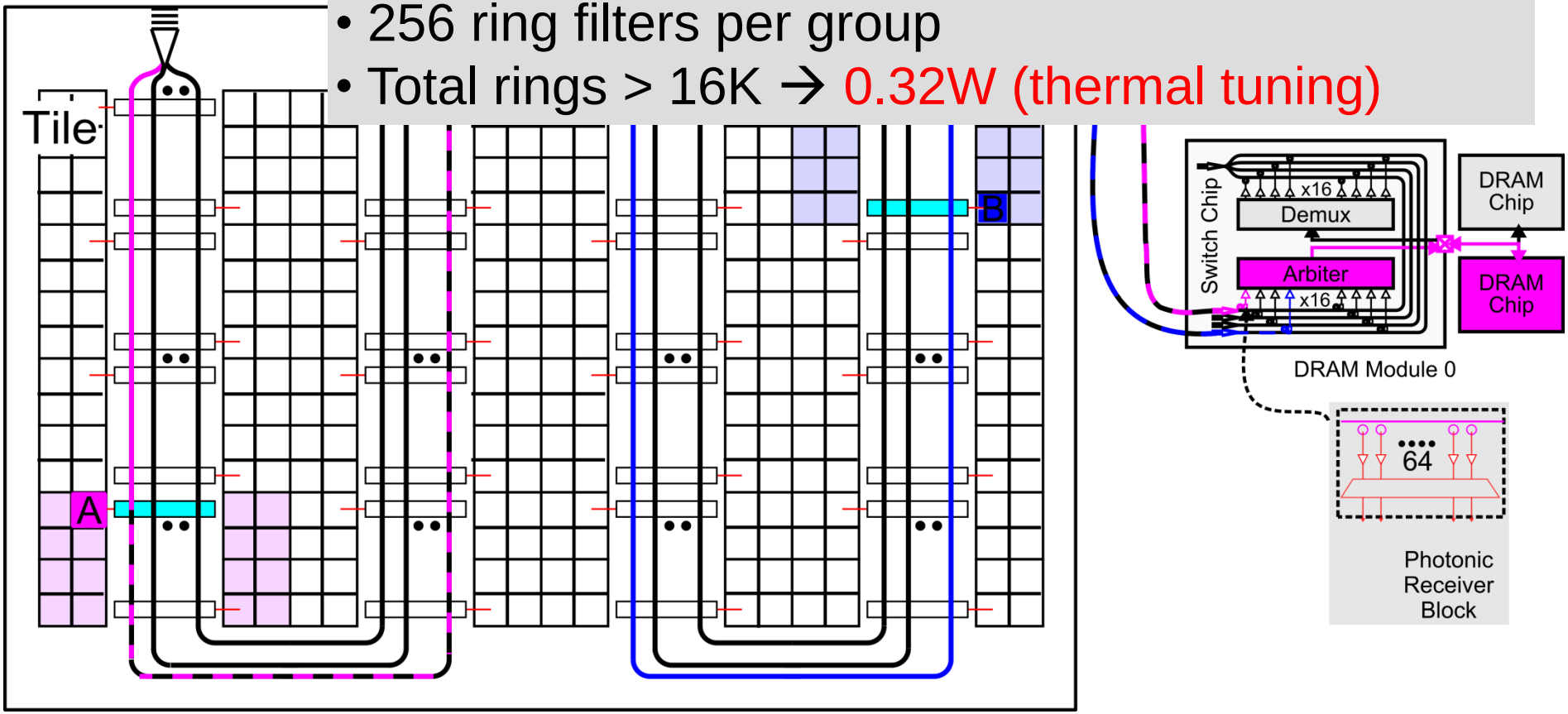
Photonic LMGS - U-shape

64-tile system w/ 16 groups, 16 DRAM Modules, 320 Gbps bi-di tile-DRAM module BW

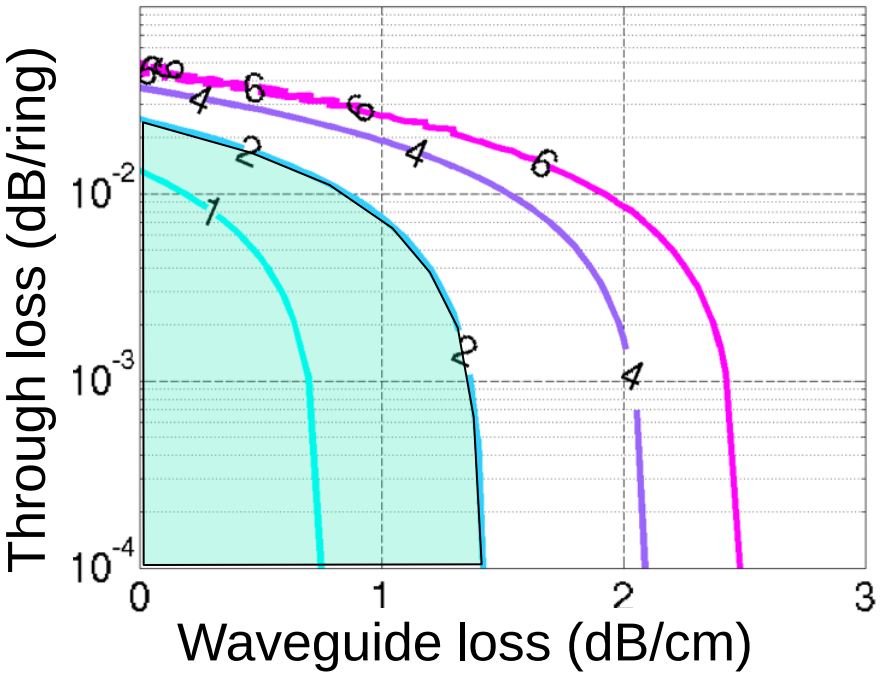


Photonic LMGS - U-shape

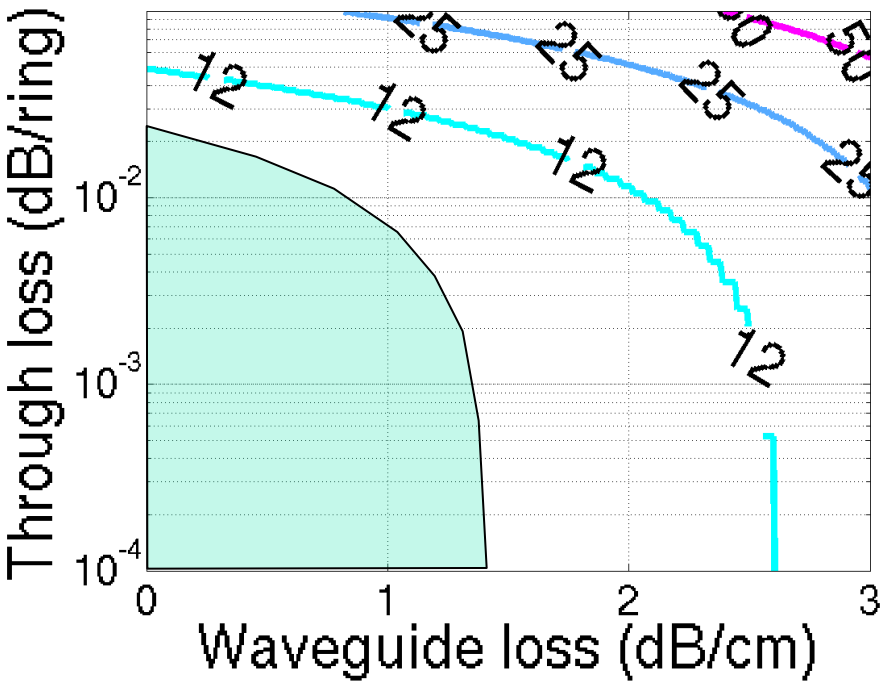
- 64 tiles
- 64 waveguides (for tile throughput = 128 b/cyc)
- 256 modulators per group
- 256 ring filters per group
- Total rings > 16K → 0.32W (thermal tuning)



Photonic device requirements in LMGS - U-shape



Optical Laser Power

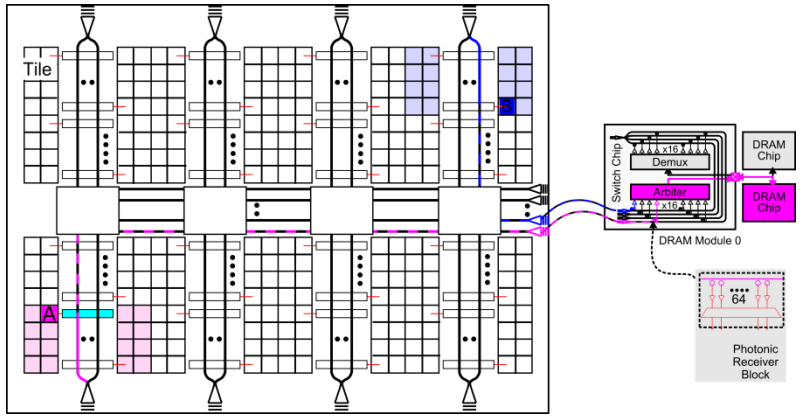


Die Area Overhead

Waveguide loss and Through loss limits for 2 W optical laser power

Photonic LMGS – ring matrix vs u-shape

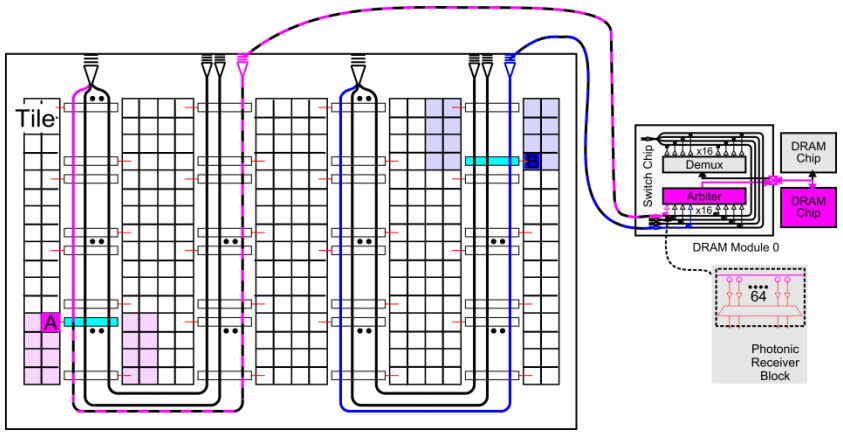
LMGS – ring matrix



- 0.64 W power for thermal tuning circuits
- 2 W optical laser power
  - Waveguide loss < 0.2 dB/cm
  - Through loss < 0.002 dB/ring

[Batten et al – Micro 2009]

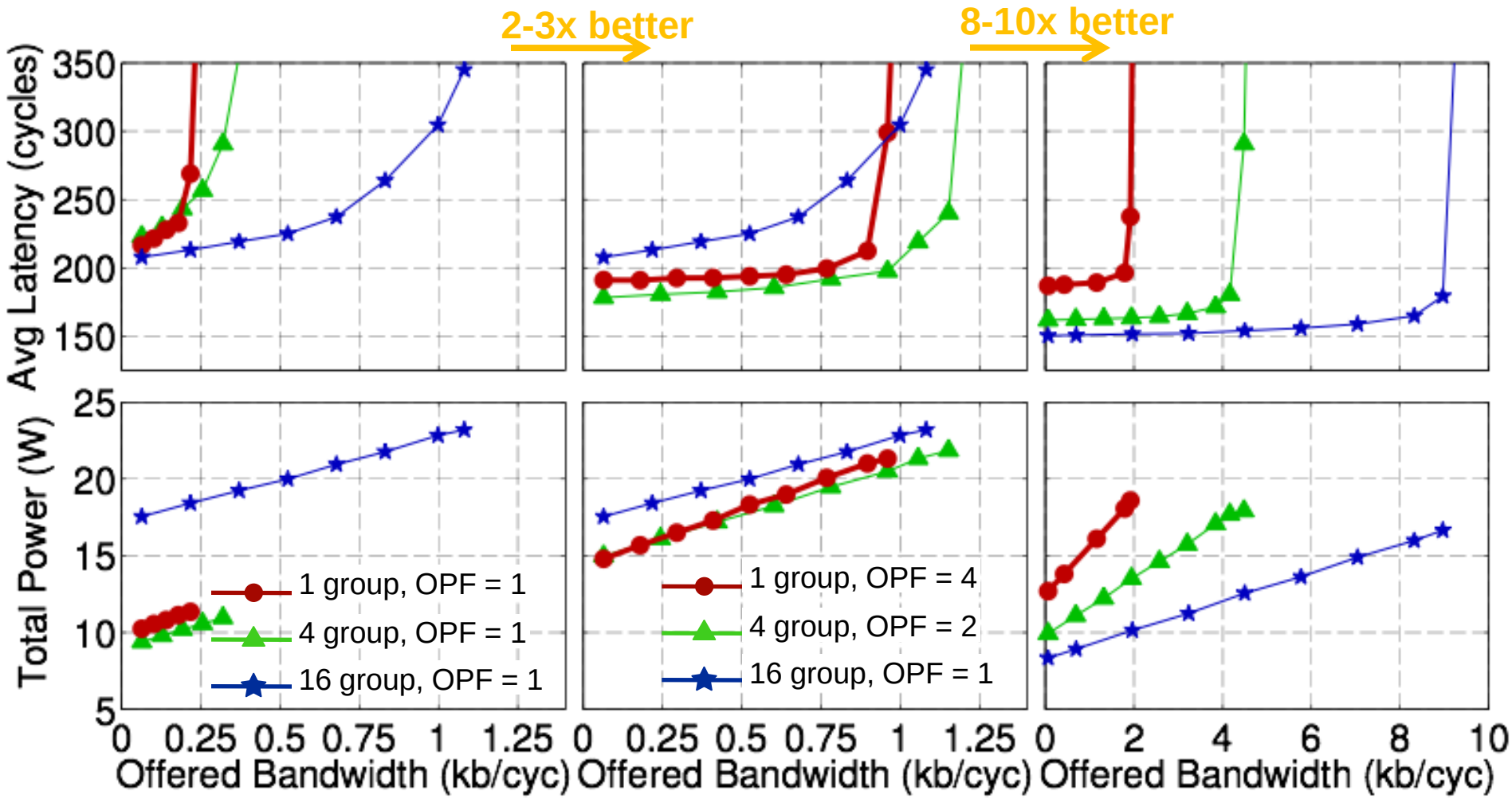
LMGS – u-shape



- 0.32 W power for thermal tuning circuits
- 2 W optical laser power
  - Waveguide loss < 1.5 dB/cm
  - Through loss < 0.02 dB/ring

[Joshi et al – PICA 2009]

Power-bandwidth tradeoff

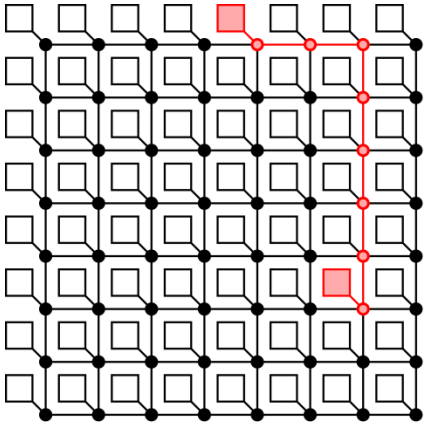


Electrical with grouping

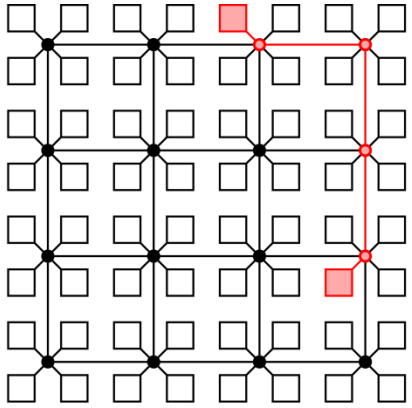
Electrical with grouping and over-provisioning

Optical with grouping and over-provisioning

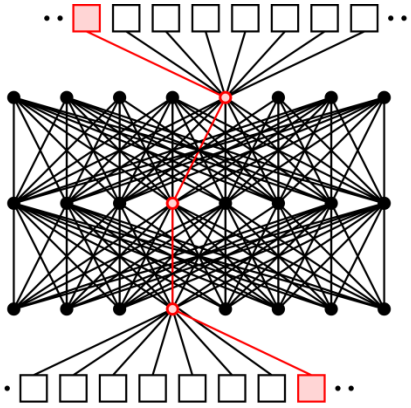
Landscape of on-chip photonic networks



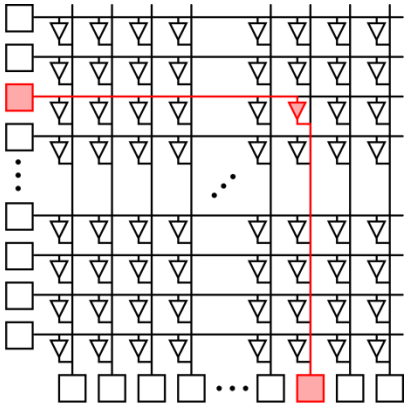
Mesh



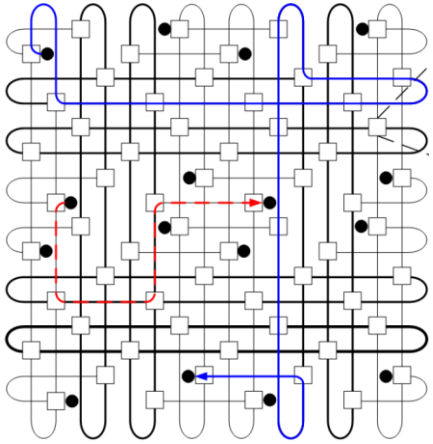
CMesh



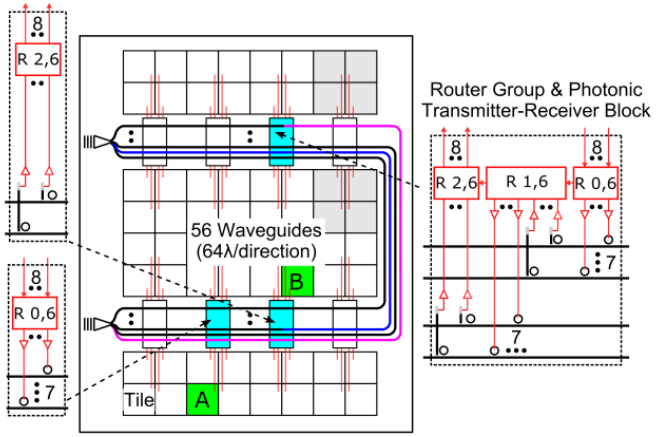
Clos



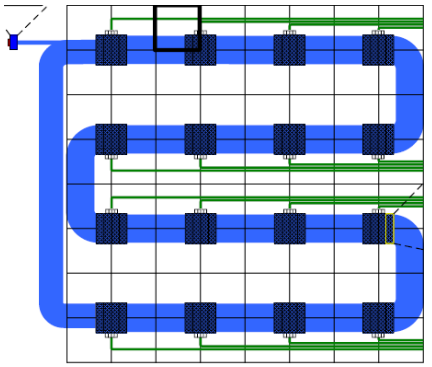
Crossbar



[Shacham'07]
[Petracca'08]



[Joshi'09a]
[Pan'09]

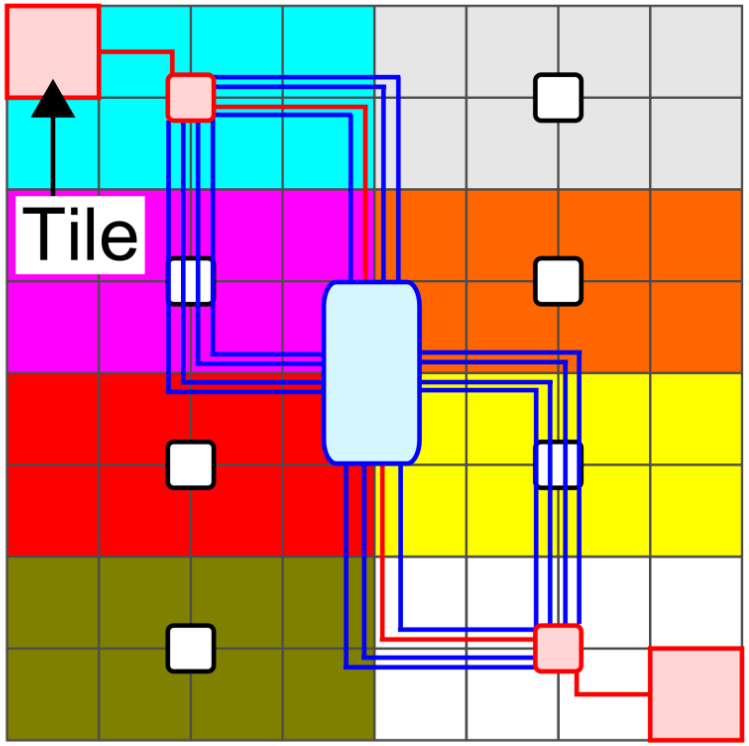


[Vantrease'08]
[Psota'07]
[Kirman'06]

Clos with electrical interconnects

8-ary 3-stage Clos

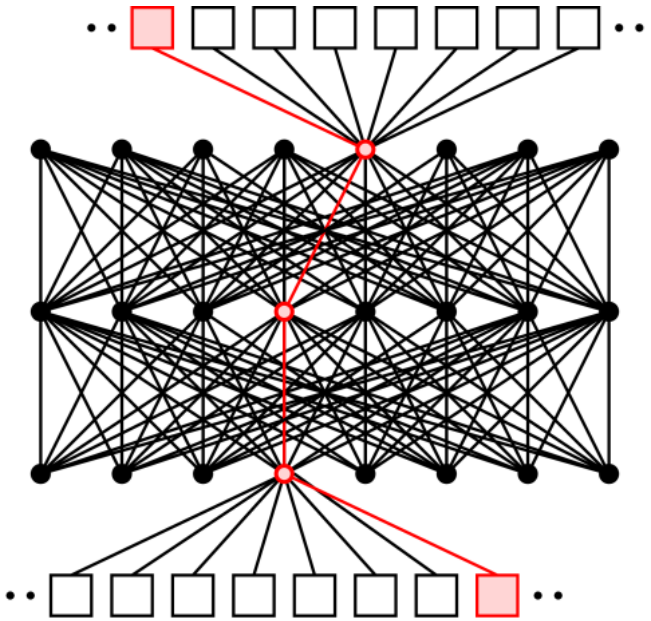
Physical mapping



Two 8 x 8 Routers

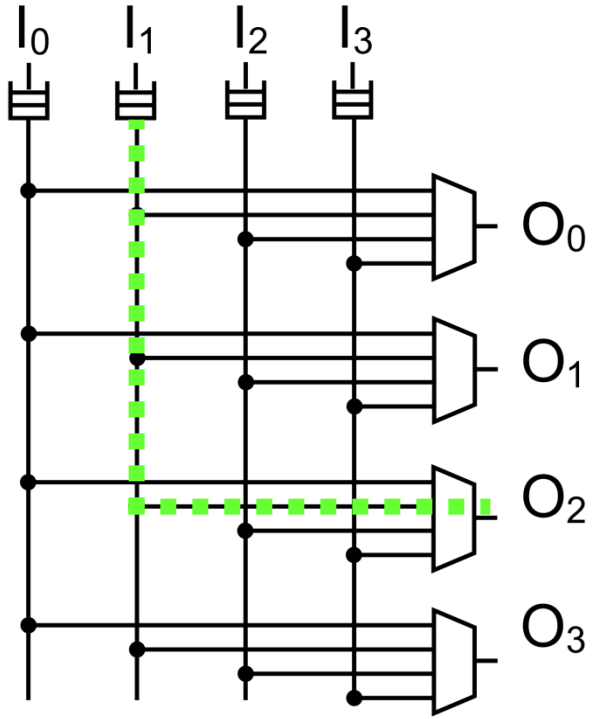
Eight 8 x 8 Routers

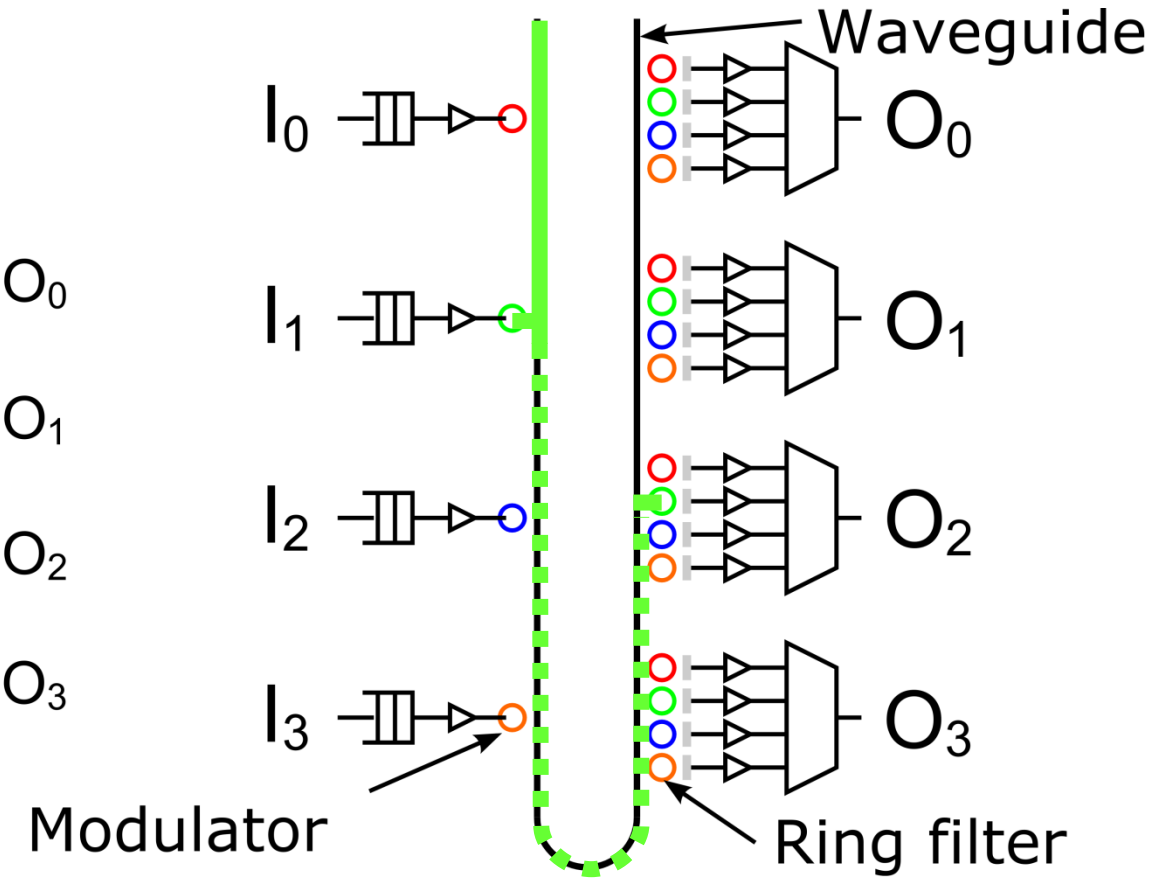
Logical topology



- 10-15 mm channels
- Pipelined Repeaters

Centralized Multiplexer Crossbar

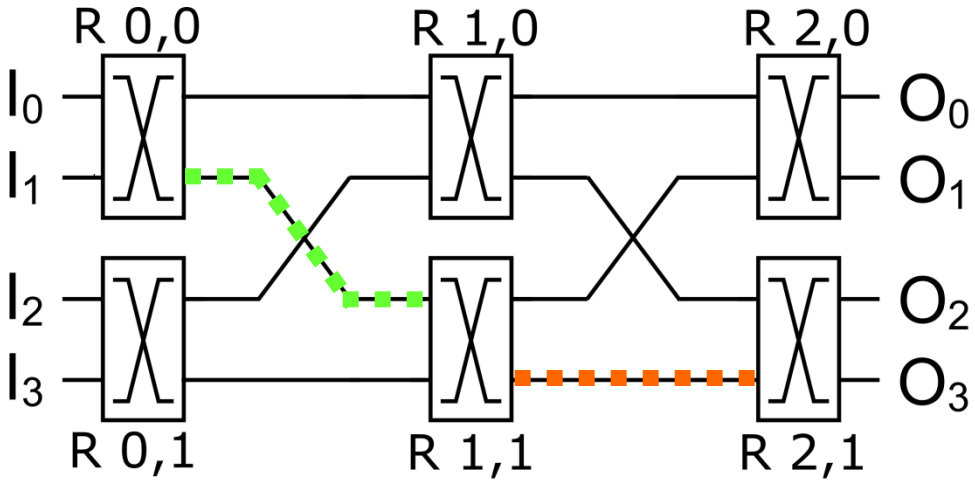




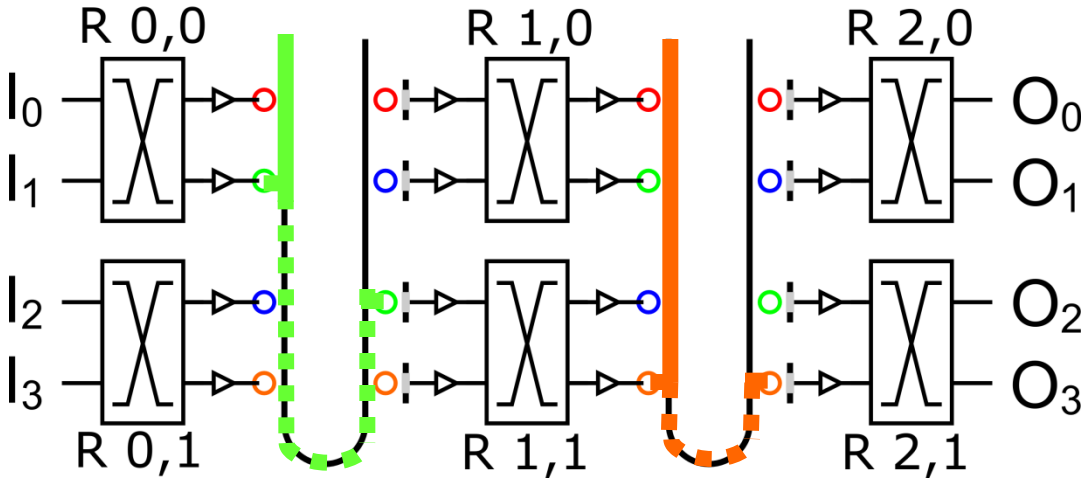
Electrical design

Photonic design

Clos network using point-to-point channels

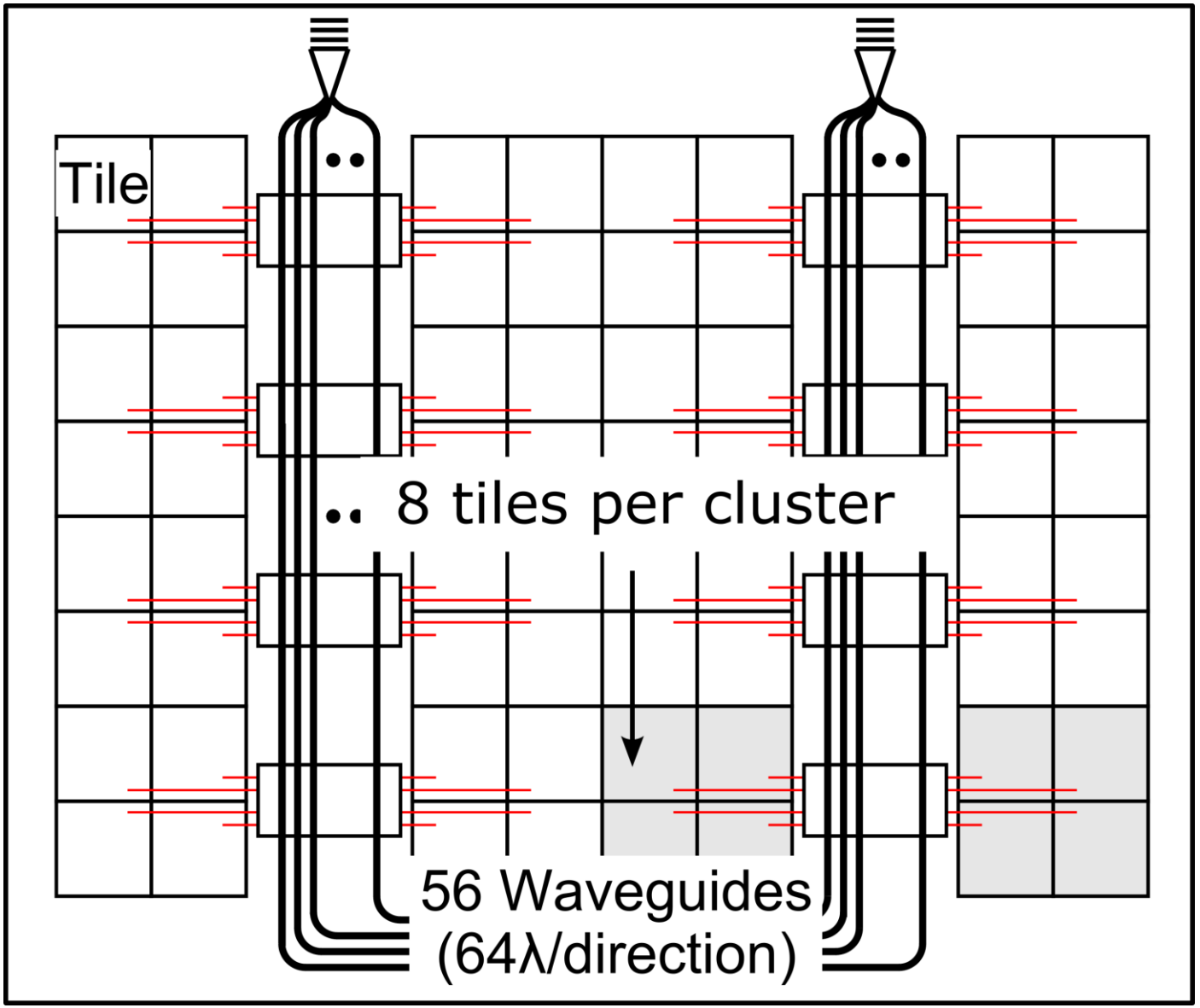


Electrical design

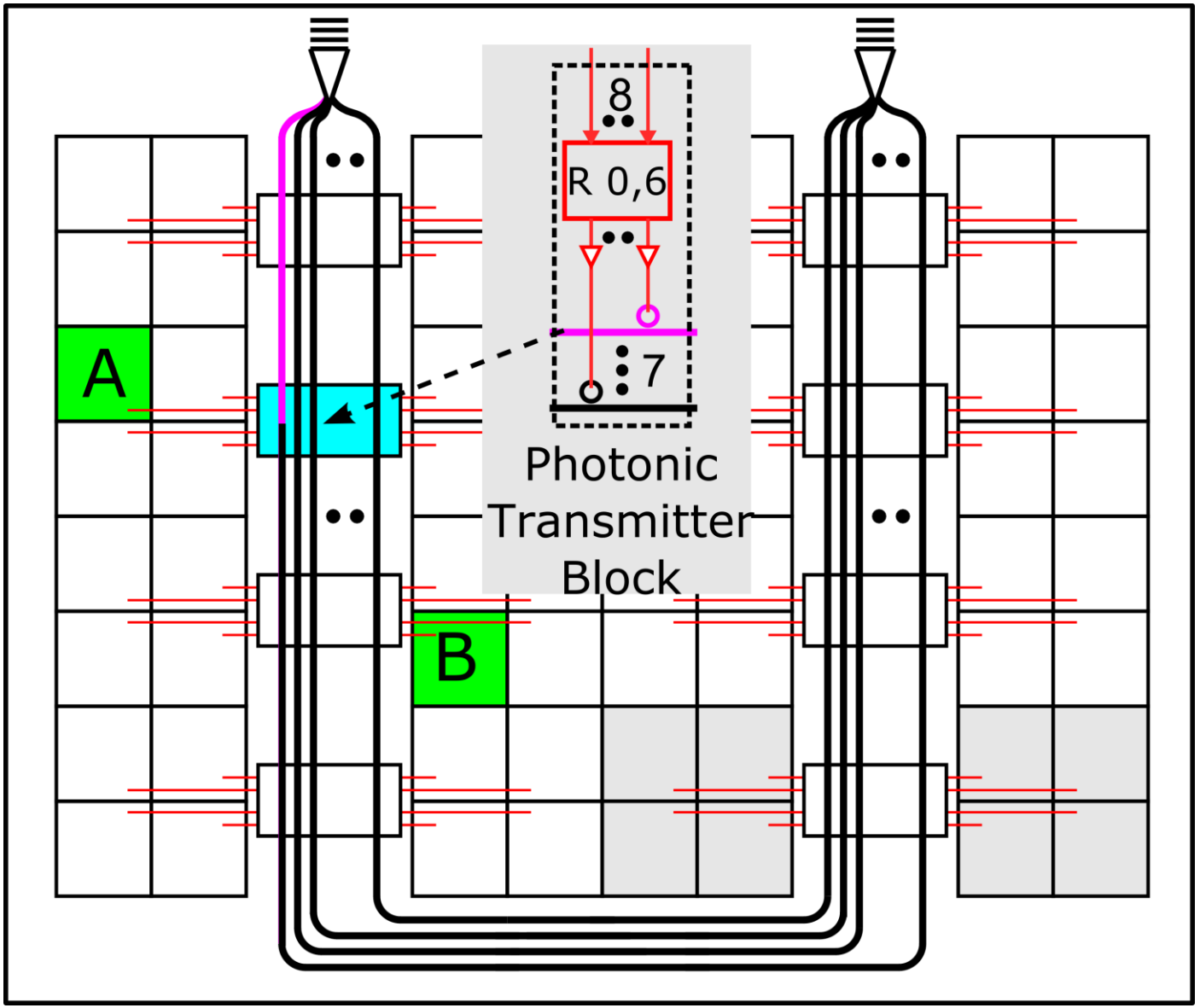


Photonic design

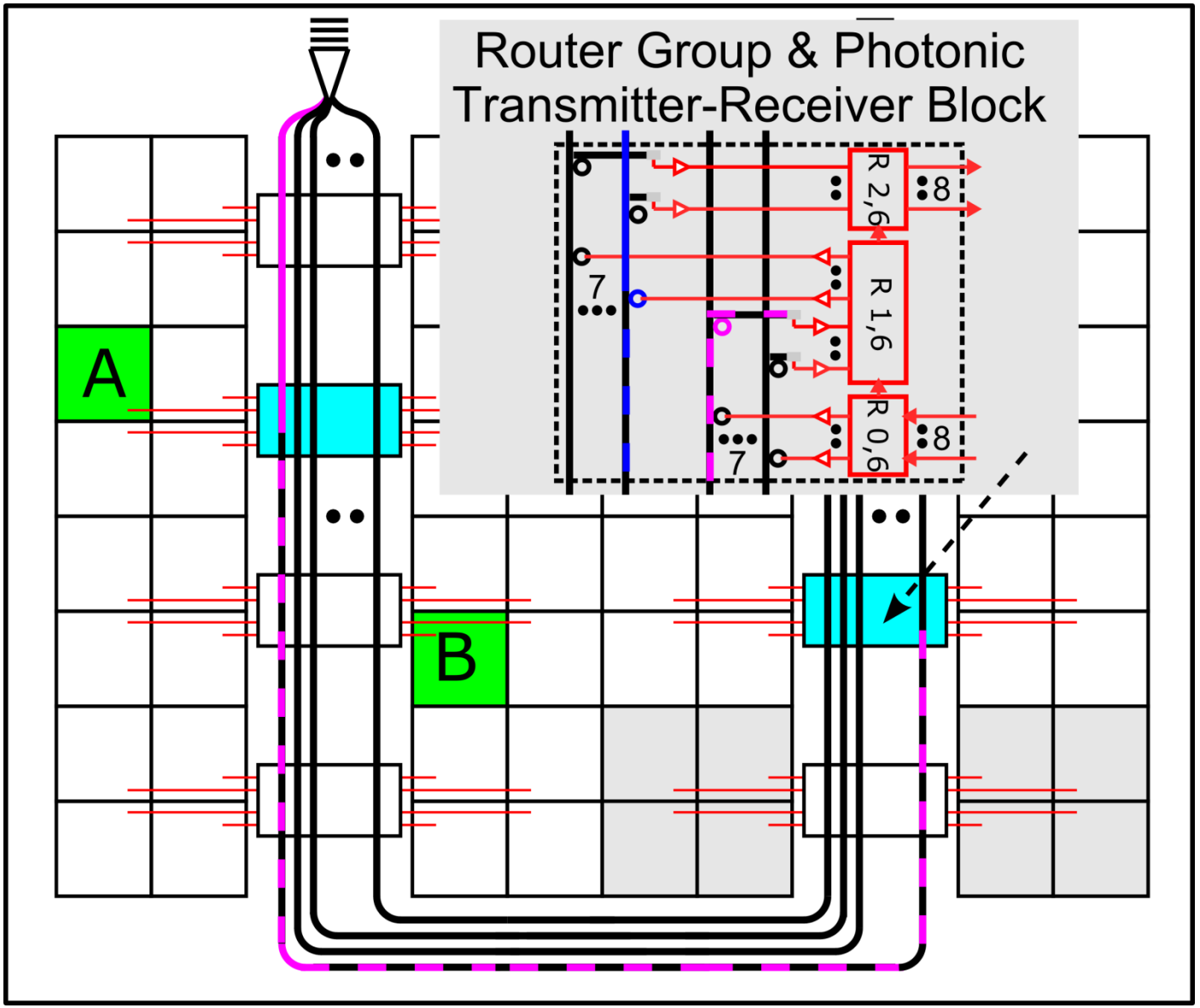
Photonic Clos for a 64-tile system



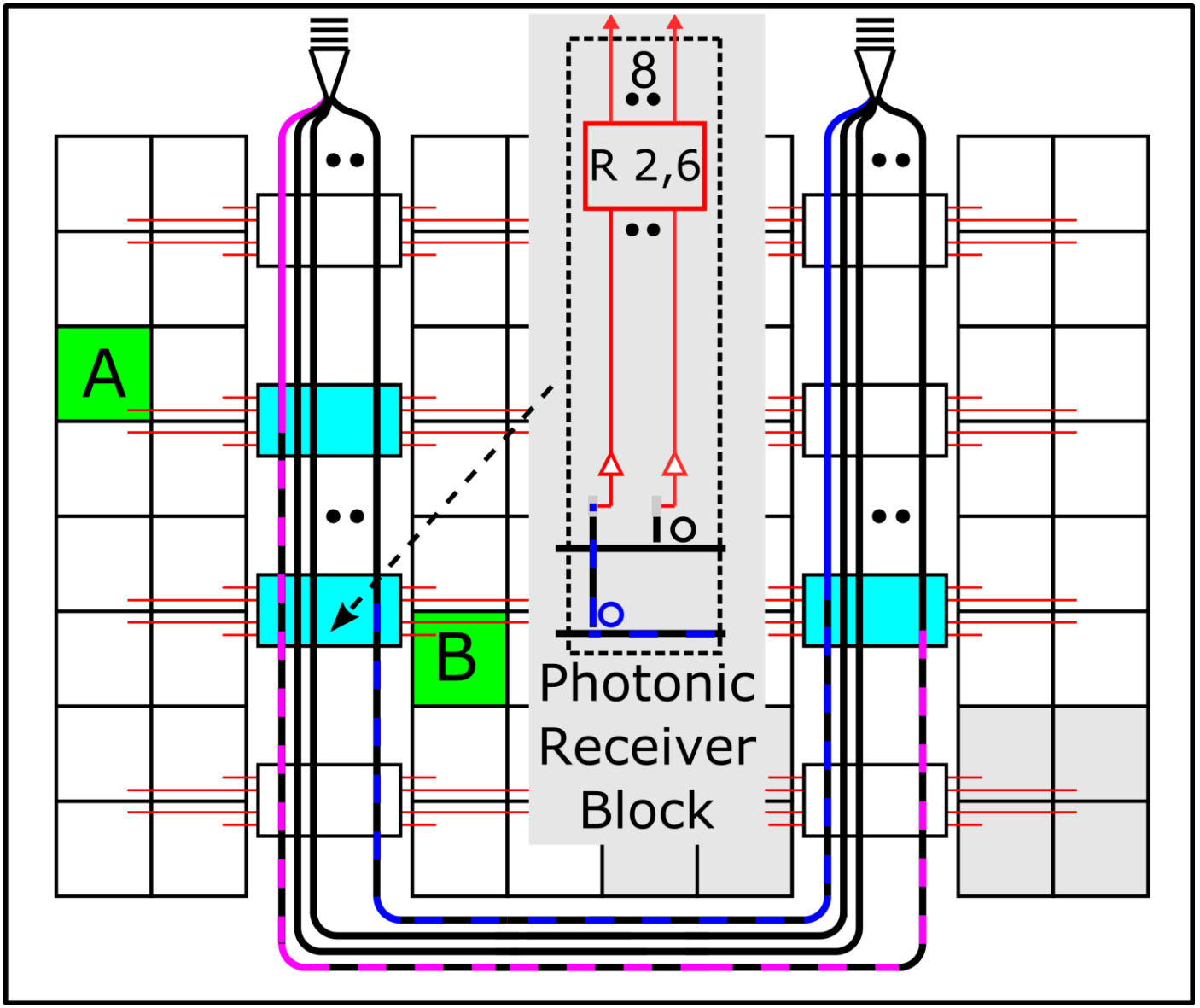
Photonic Clos for a 64-tile system



Photonic Clos for a 64-tile system

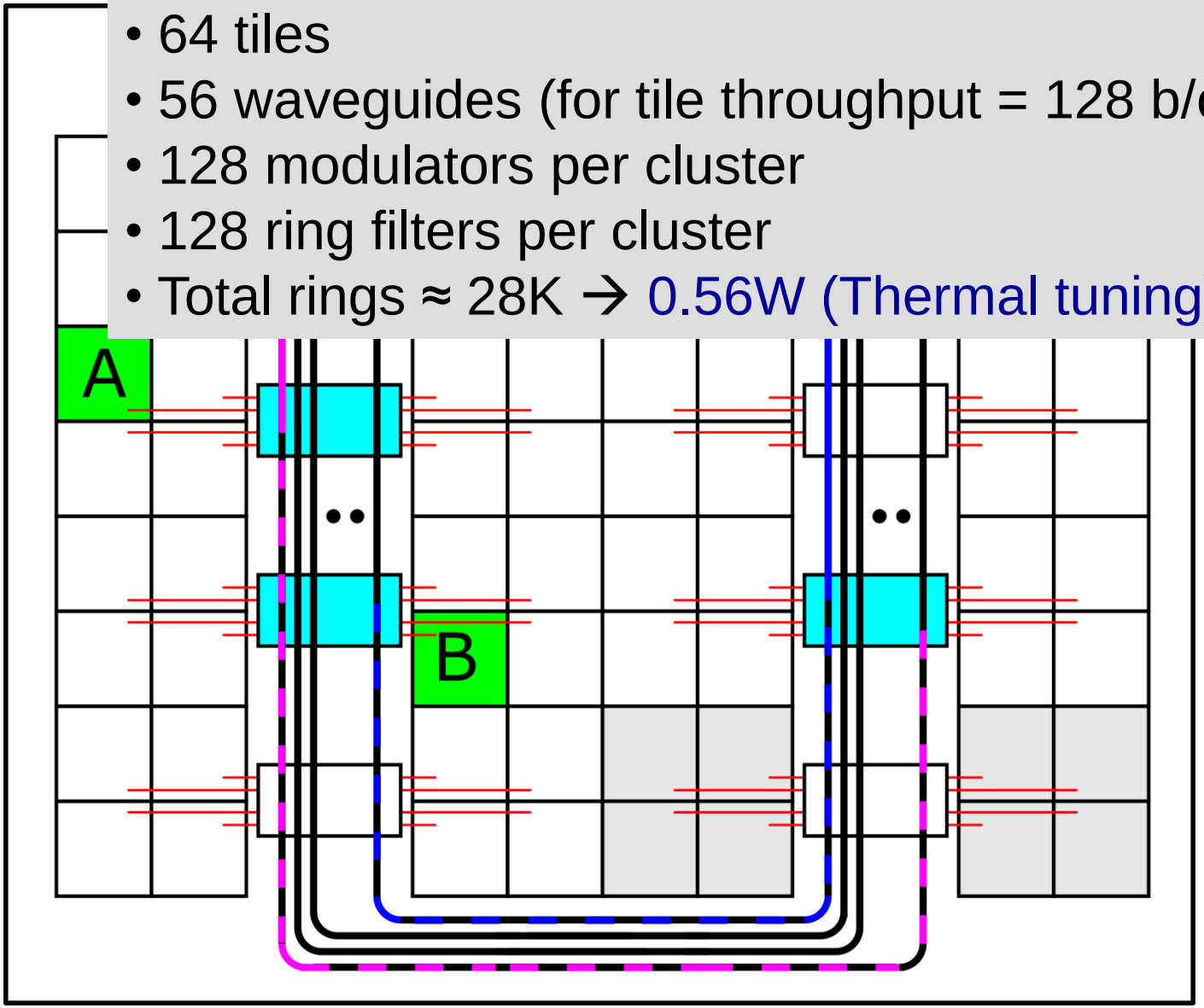


Photonic Clos for a 64-tile system

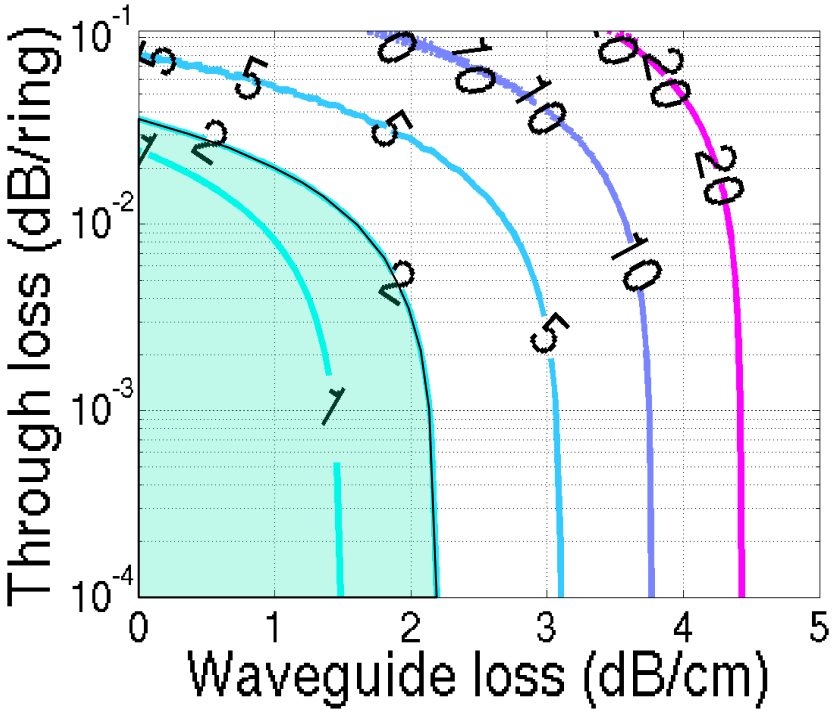


Photonic Clos for a 64-tile system

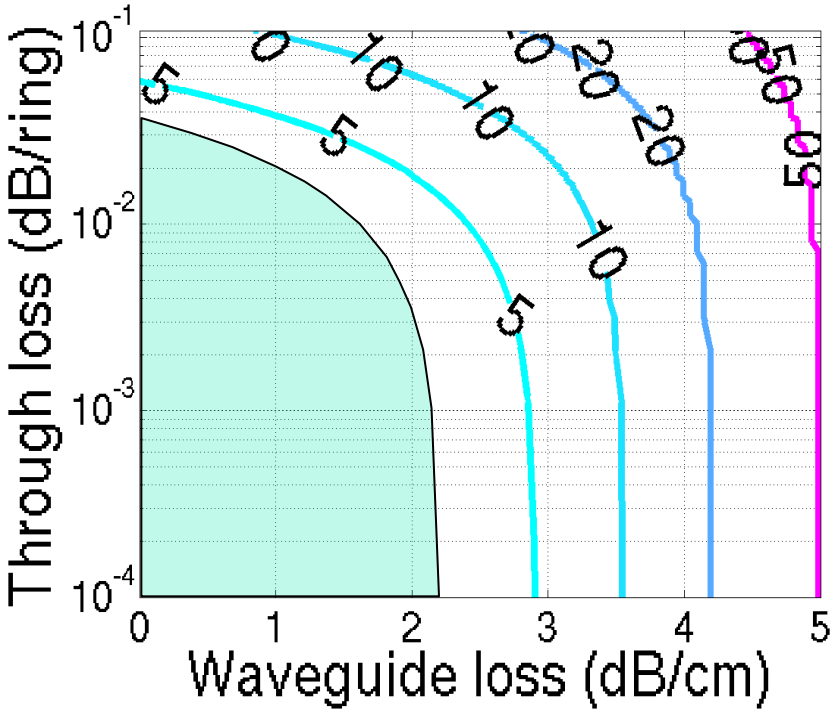
- 64 tiles
- 56 waveguides (for tile throughput = 128 b/cyc)
- 128 modulators per cluster
- 128 ring filters per cluster
- Total rings ≈ 28K → 0.56W (Thermal tuning)



Photonic device requirements in a Clos



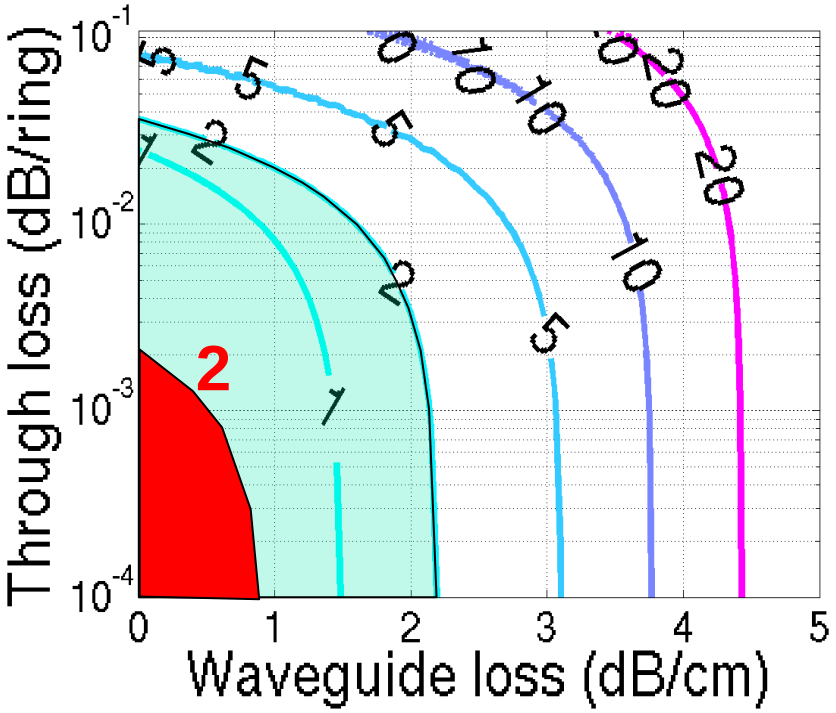
Optical laser power (W)



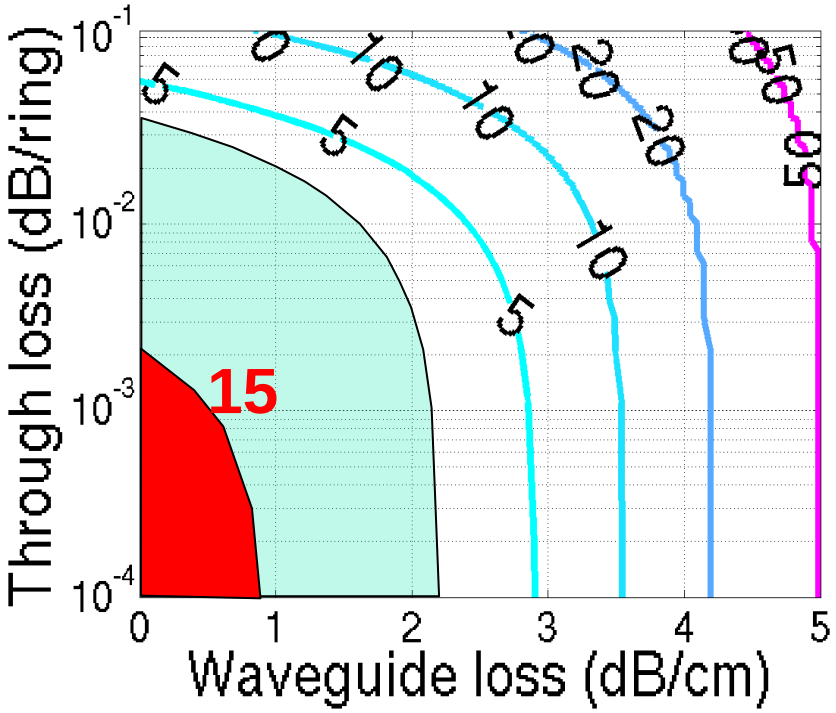
Percent die area for photonic devices

Waveguide loss and Through loss limits for 2 W optical laser power constraint

Photonic device requirements in a Clos



Optical laser power (W)



Percent die area for photonic devices

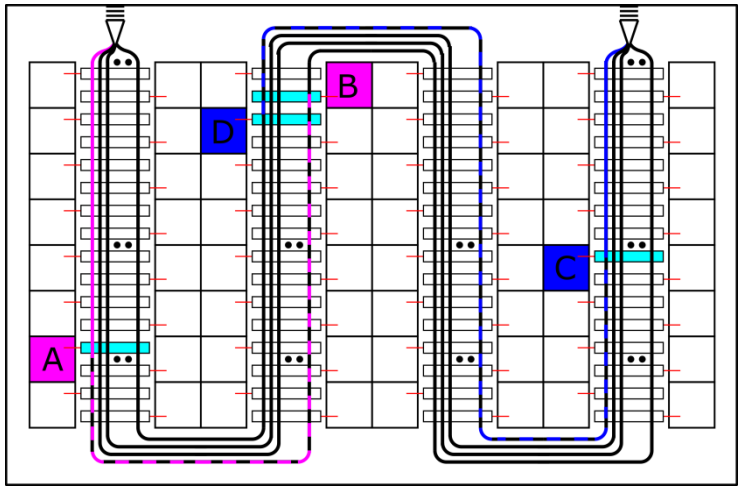
Optical loss tolerance for Crossbar

Optical loss tolerance for Clos

2 W optical power contours

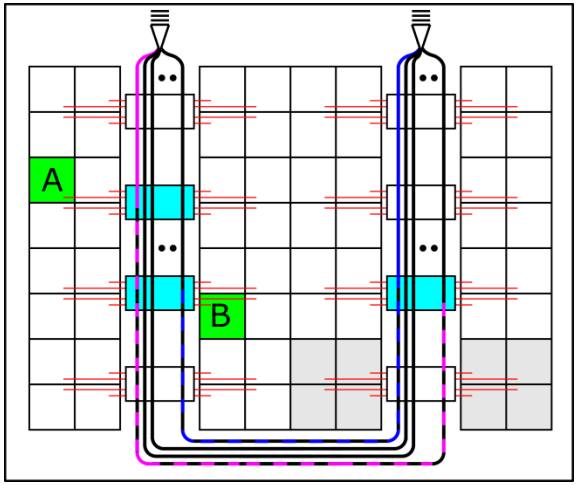
Photonic Crossbar vs Photonic Clos

Crossbar



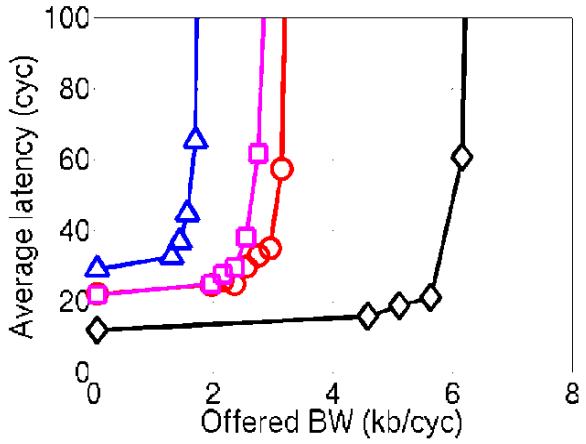
- 10 W power for thermal tuning circuits
- For 2 W optical laser power
  - Waveguide loss < 1 dB/cm
  - Through loss < 0.002 dB/ring

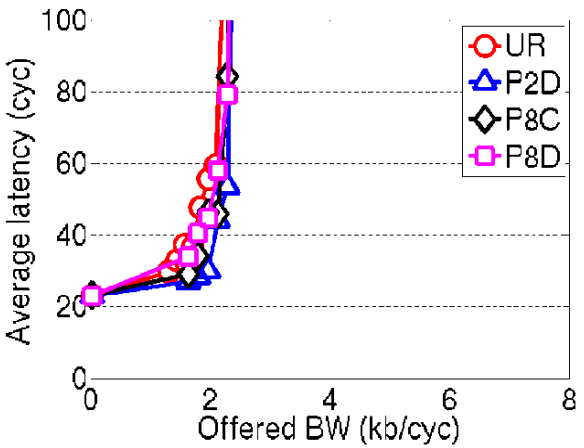
Clos

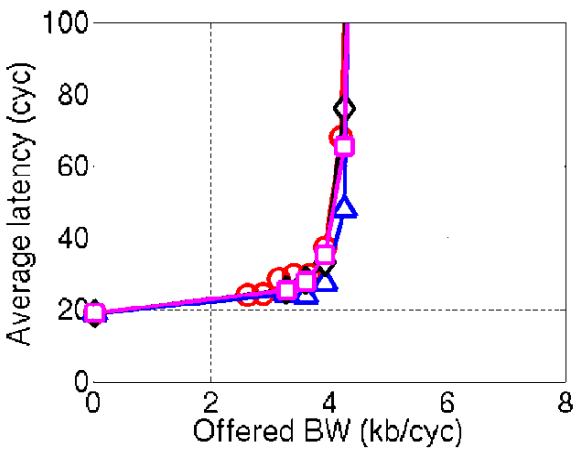


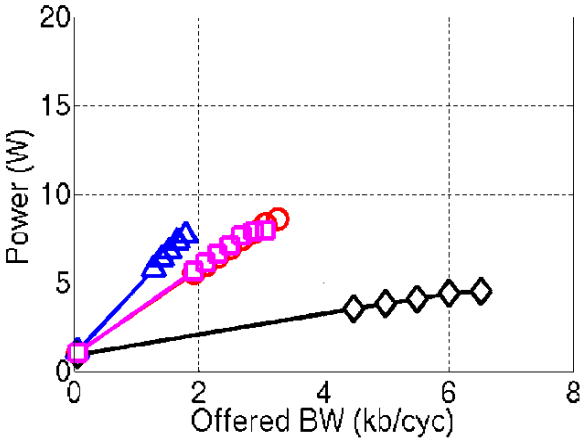
- 0.56 W power for thermal tuning circuits
- For 2 W optical laser power
  - Waveguide loss < 2dB/cm
  - Through loss < 0.05 dB/ring

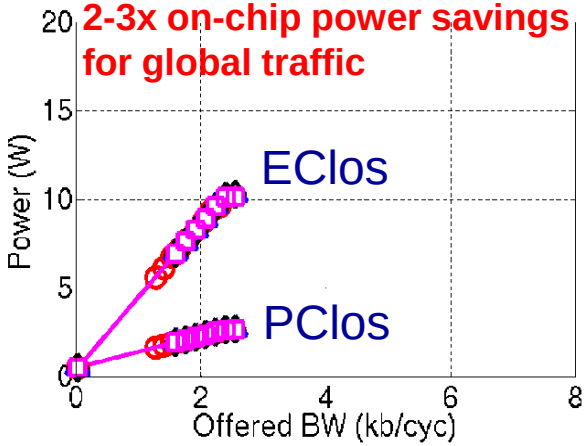
Power-Bandwidth tradeoff

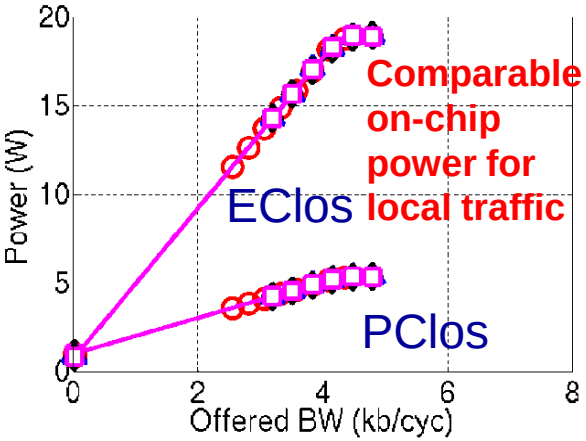












	CMeshX2	Clos	Clos
b_c	128b	64b	128b
θ_T	4kb/cycle	4kb/cycle	8kb/cycle

Conclusion

- Computer interconnects are very complex micro-communication systems
- Cross-layer design approach is needed to solve the on-chip and off-chip interconnect problem
  - Most important metrics
    - Bandwidth-density (Gb/s/um)
    - Energy-efficiency (mW/Gb/s)
  - Monolithic CMOS-photonics can improve the throughput by 10-20x
  - But, need to be careful
    - Optimize network design (electrical switching, optical transport)
    - Use aggregation to increase link utilizations
    - Optimize physical mapping (layout) for low optical insertion loss