Scalable Nanophotonic Interconnect for Cache Coherent Multicores

Randy W. Morris, Jr. and Avinash K. Kodi
Department of Electrical Engineering and Computer Science
Ohio University, Athens, OH 45701
E-mail: rmorris@cs.ohiou.edu, kodi@ohio.edu
Website: http://oucsace.cs.ohiou.edu/~avinashk/

WINDS 2010: Workshop on the Interaction between Nanophotonic Devices and Systems
Atlanta, GA
December 5, 2010

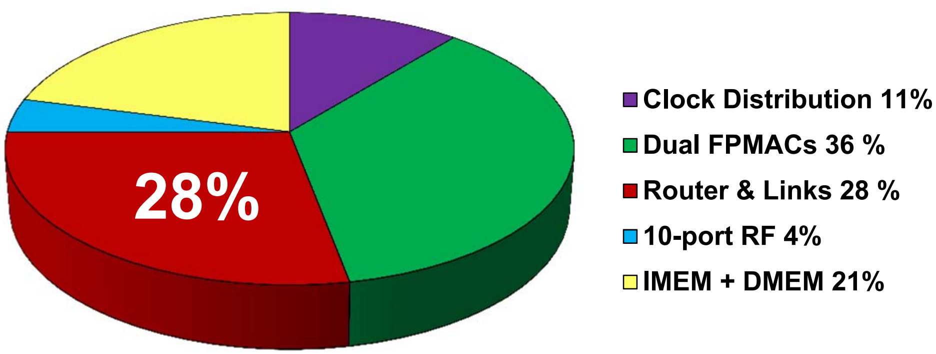
Talk Outline

- **Section I: Motivation & Background**
- **Section II: Dual Sub-Network for Snoopy Cache Coherent Nanophotonic Architecture**
- **Section IV: Performance Analysis**
- **Section V: Future Work**

Why Nanophotonics?

- Power consumption of Network-on-Chips (NoCs) [1] using metallic interconnects is projected to exceed expectation by a **factor of 10**

Tile Power: Intel Tera-Flops (65 nm)[2]



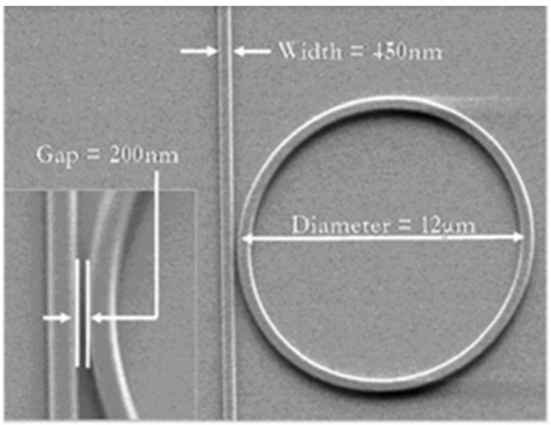
Nanophotonic Technology

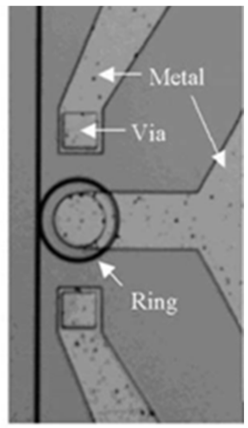
- **Low Power**
- **Small Footprint (10 – 15 μm)**
- **High Bandwidth (10 – 20 Gbps)**
- **CMOS Compatibility**

1. Reference : J.D.Owens, W.J.Dally, R.Ho, D.N.Jayasimha, S.W.Keckler and L.S.Peh, “Research Challenges for On-Chip Interconnection Networks”, IEEE Micro, vol. 27, no. 5, pp. 96 – 108, September-October 2007.

2. Y. Hoskote, “A 5-GHz Mesh Interconnect for A Teraflops Processor,” IEEE Computer Society, 2007 pp. 51-61

Micro-ring Resonators





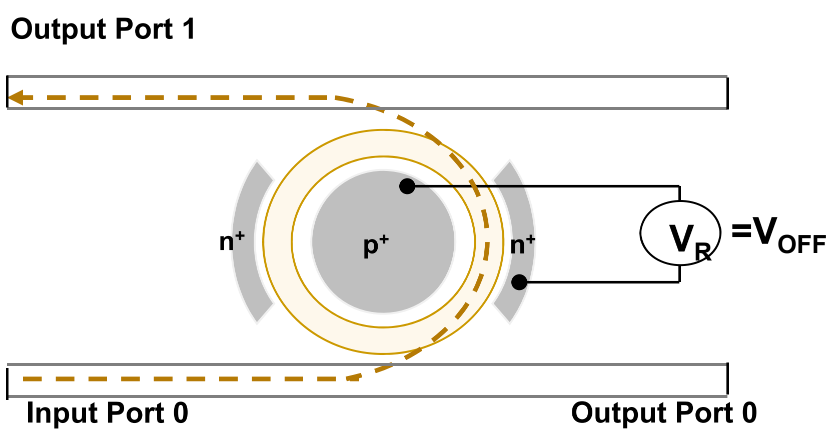
Resonant wavelength (λ_0)

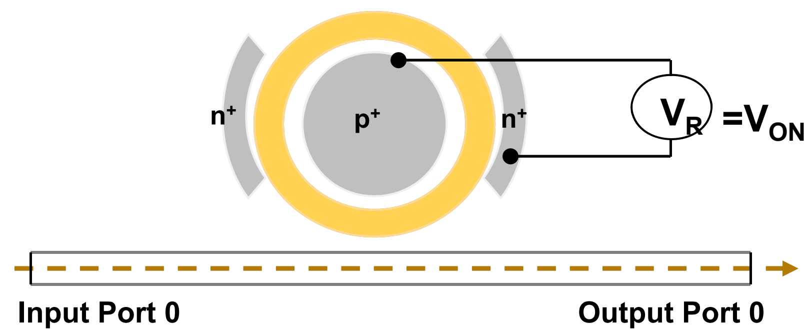
$$\boxed{\lambda_0 \times m = n_{eff} \times 2\pi R}$$

$m \rightarrow$ an integer

$n_{eff} \rightarrow$ effective refractive index

$R \rightarrow$ radius of the ring resonator



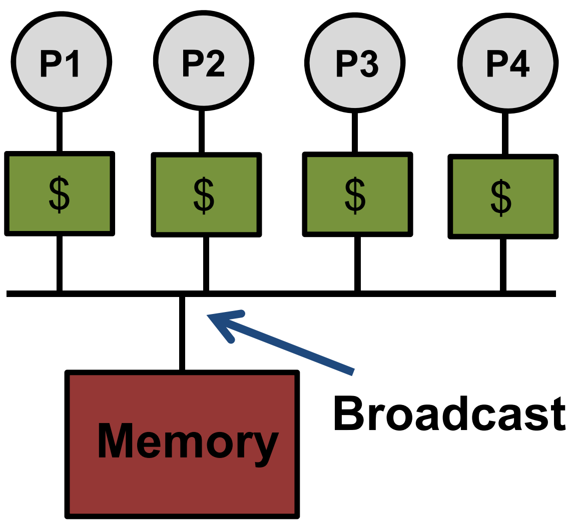


1. Lipson, M., Compact Electro-Optic Modulators on a Silicon Chip, IEEE J. Sel. Top. Quant., Vol. 12, No. 6, Nov.-Dec. 2006, p. 1520-6.
2. M. Lipson, Guiding, Modulating and Emitting Light on Silicon - Challenges and Opportunities, IEEE Journal of Lightwave Technologies, Vol. 23, No. 12, 12 December 2005 (invited).

Cache Coherence

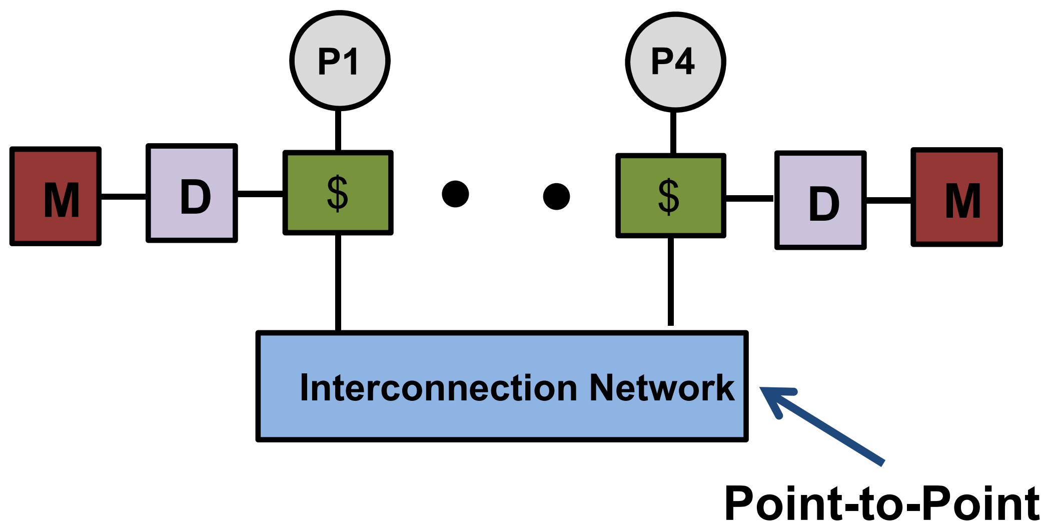
- **Write propagation** (write by any processor should become visible to all other processors)
- **Write serialization** (all writes from same or different processors are seen in the same order by all processors)

Snoopy Protocols



Easy to Program
Not Easily Scalable

Directory Protocols



Scalable
High miss latency

Problems with Snoopy Networks

Two major problems with snoopy cache coherent networks

(1) Interconnect bandwidth for broadcasting of memory requests

- Bus Networks: Limits one request per cycle
- Multiple Buses: Increases cache controllers
- Point-to-Point Networks: Selective multicasting & Ordering

(2) Cache Access Rate

- Cache tag lookup (latency)
- Increased power consumption

Related Work (to name a few)

Electrical

- Split Transactional Bus
- Sun Fireplane (SC 2001)
- Timestamp Snooping (ASPLOS 2000), Multicast Snooping (ISCA 2001
- Jetty (HPCA 2001), Region Scout (ISCA 2005), Intel QPI
- Broadcasting on Ordered Networks (HPCA 2009, MICRO 2009)

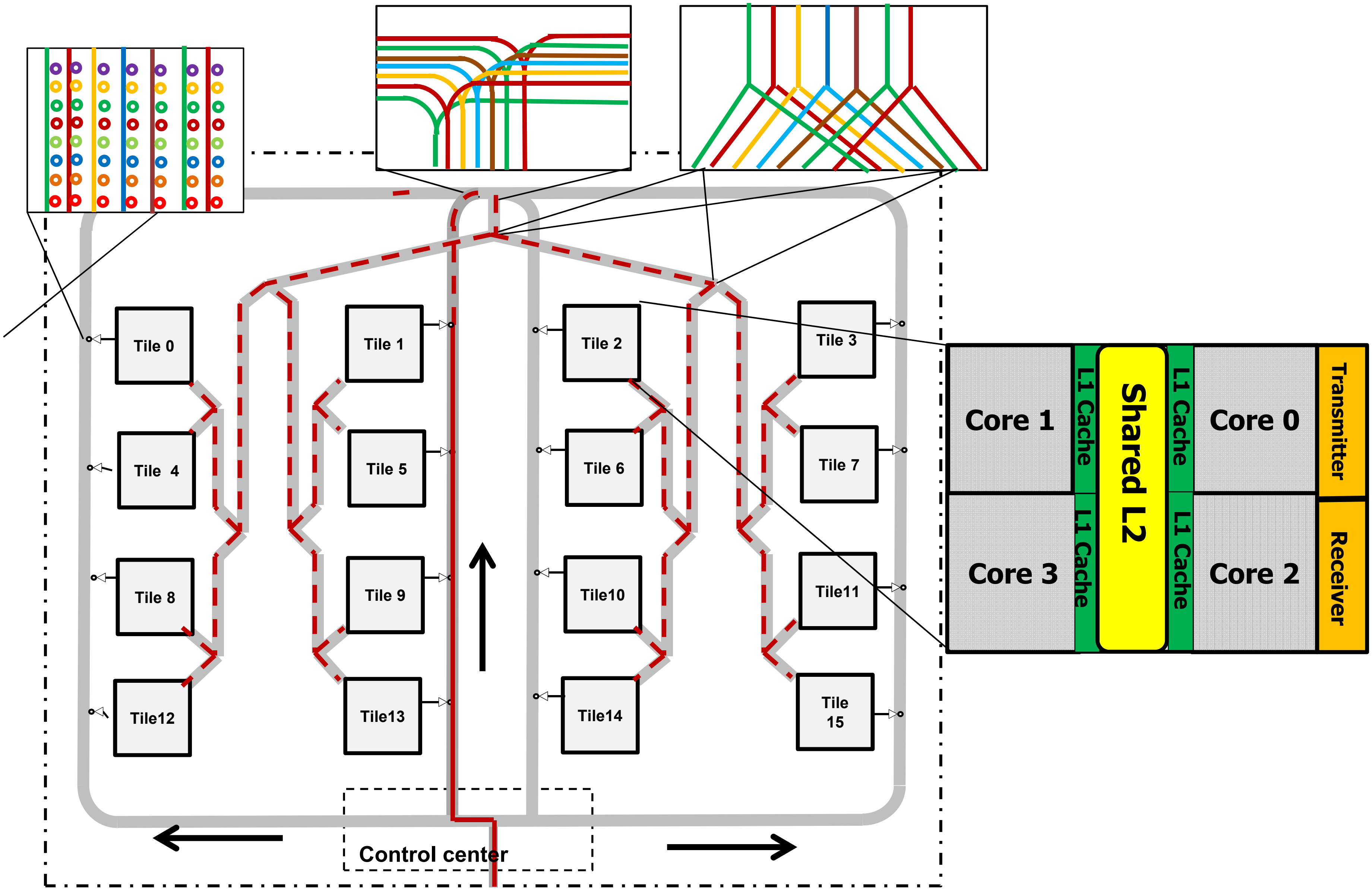
Optical/Nanophotonic

- SYMNET (Trans on Parallel & Dist Systems 2004)
- Shared Bus (MICRO 2006), Wavelength Routed Oblivious Network (ASPLOS 2010)
- Spectra (ISPLED 2009), ATAC (PACT 2010)

CC-NPA Architecture

- **Advantages of the proposed architecture**
  - Dual sub-networks for memory request
    - Broadcast & Multicast networks
  - Broadcast network used by all tiles to fetch the missed block
    - Network access implemented using tokens
    - Determines the sharing pattern
  - Multicast network to be shared between nodes to send selective requests
    - Reduces the broadcast requirement
    - Simultaneous transient requests in progress to different memory locations
  - Reducing the external laser power by unique power guiding techniques

Proposed Broadcast Sub-Network Architecture: CC-NPA

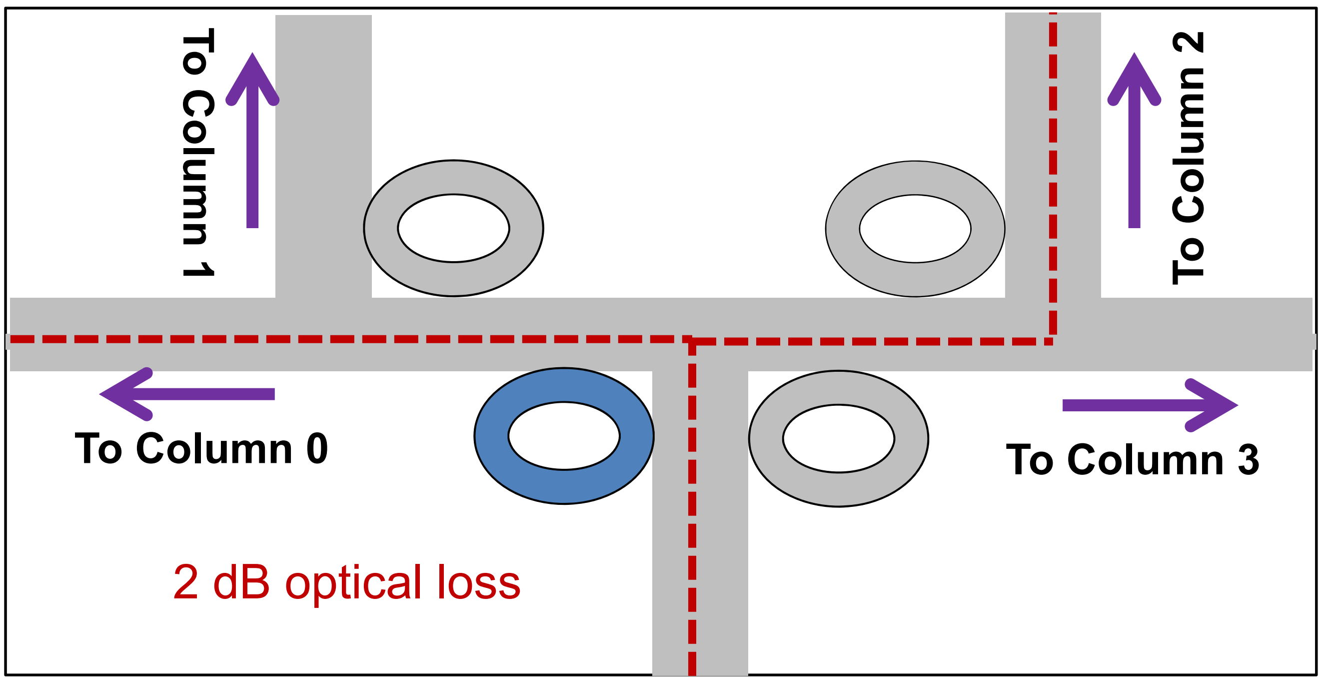


Power Guiding

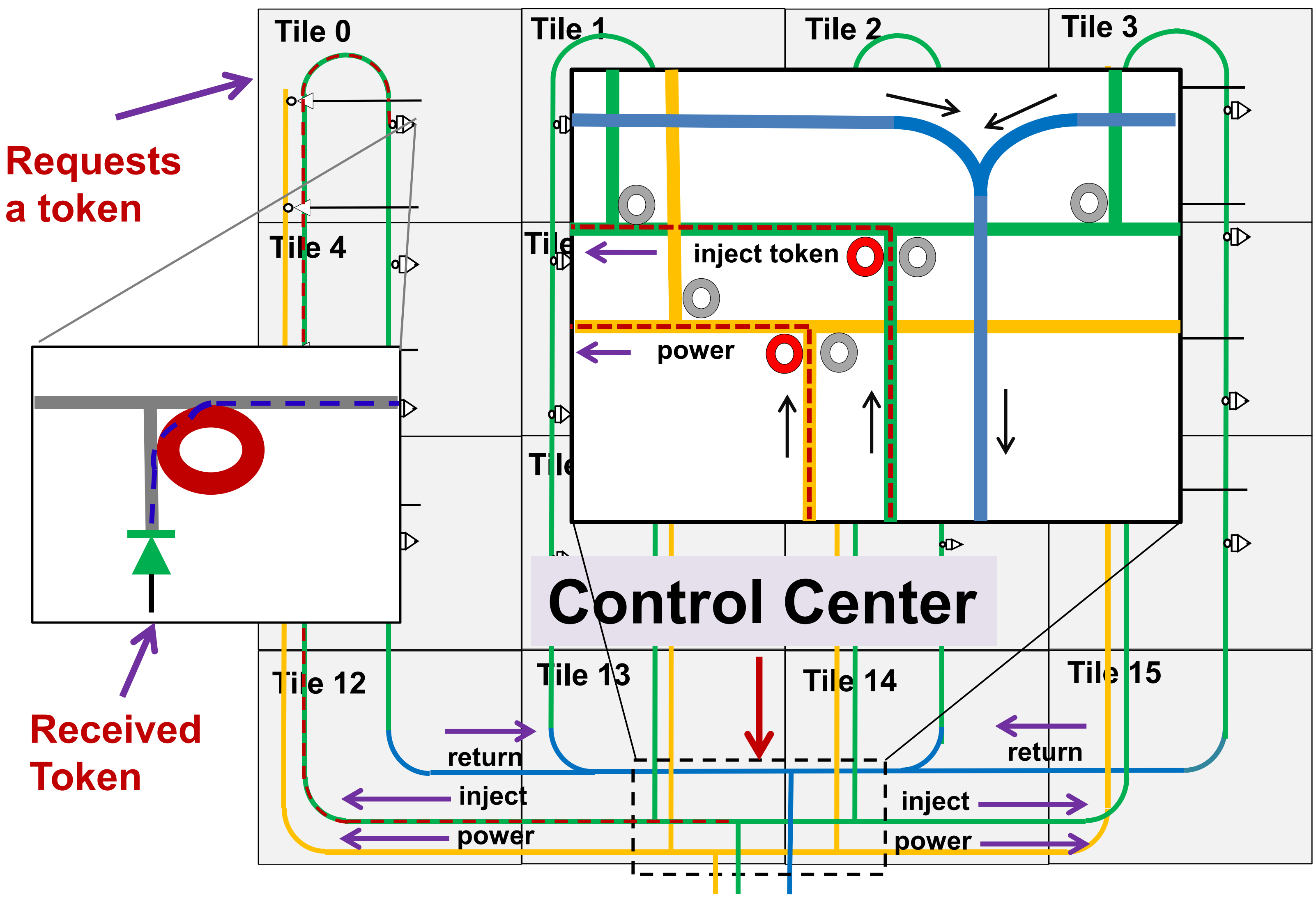
As only one core can transmit, route power to a column of cores.

- Reduction in optical power (~75%)

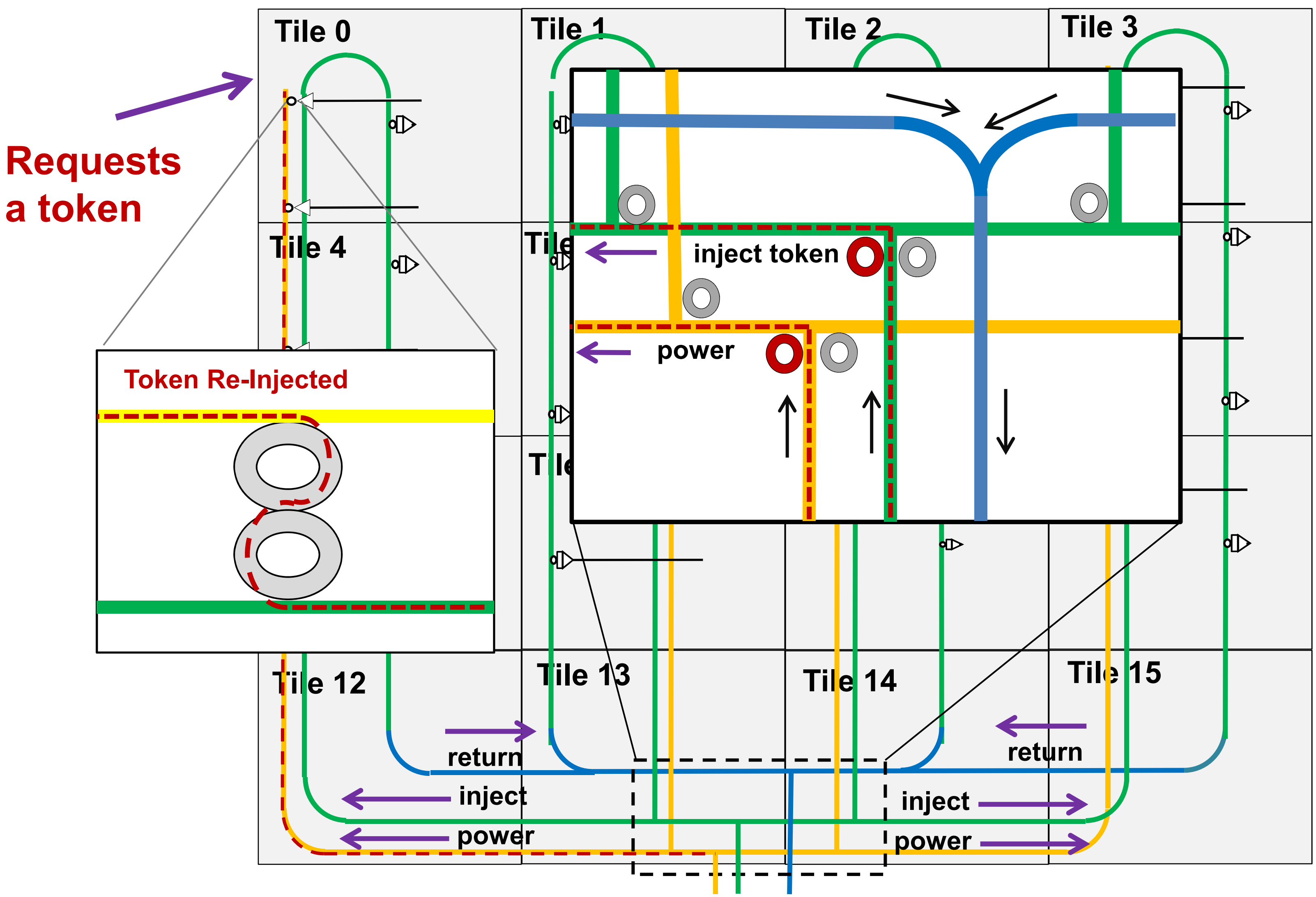
The active column is determined by the circulating optical tokes



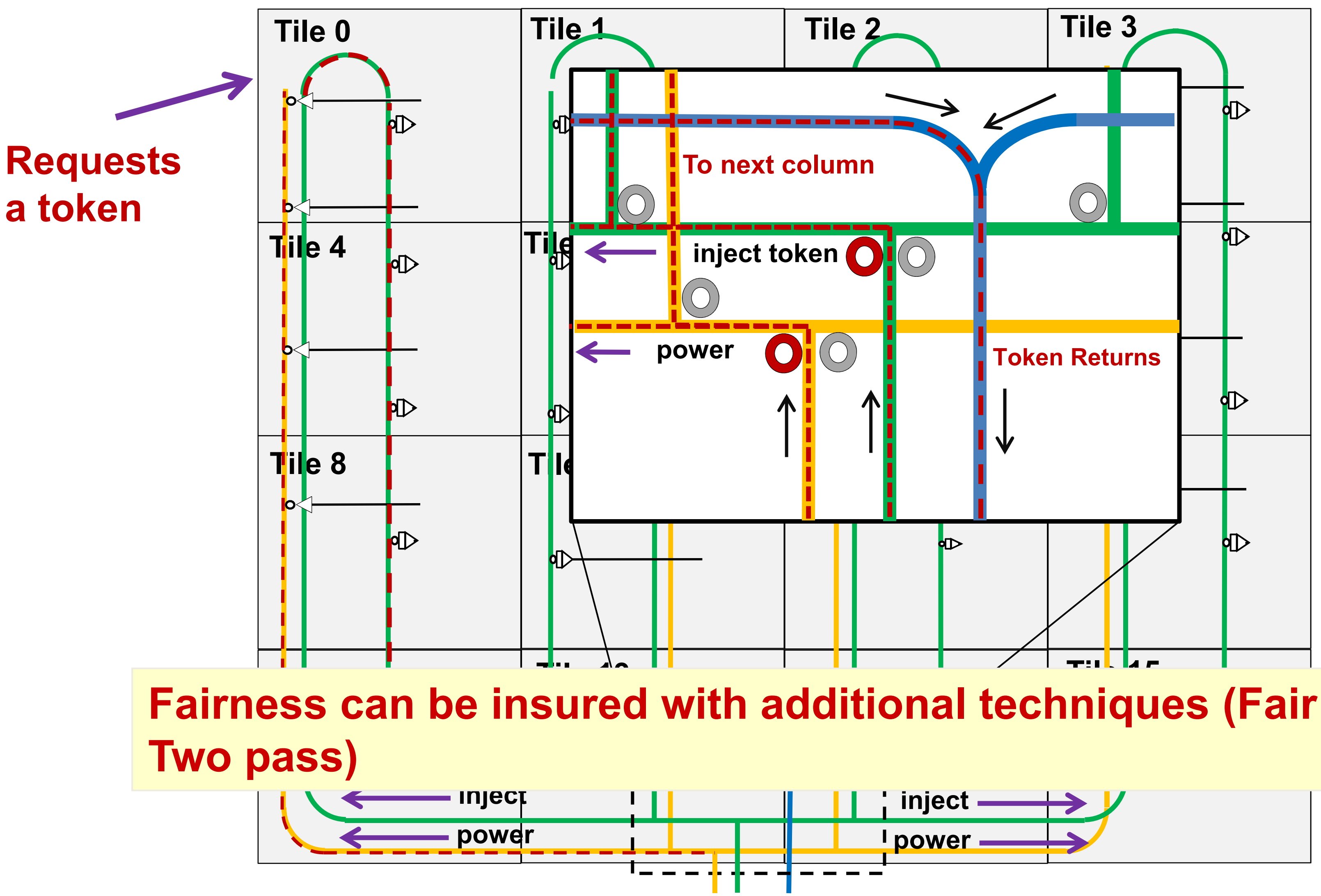
Optical Token System (1/3)



Optical Token System (2/3)



Optical Token System (3/3)



Fairness can be insured with additional techniques (Fair slot, Two pass)

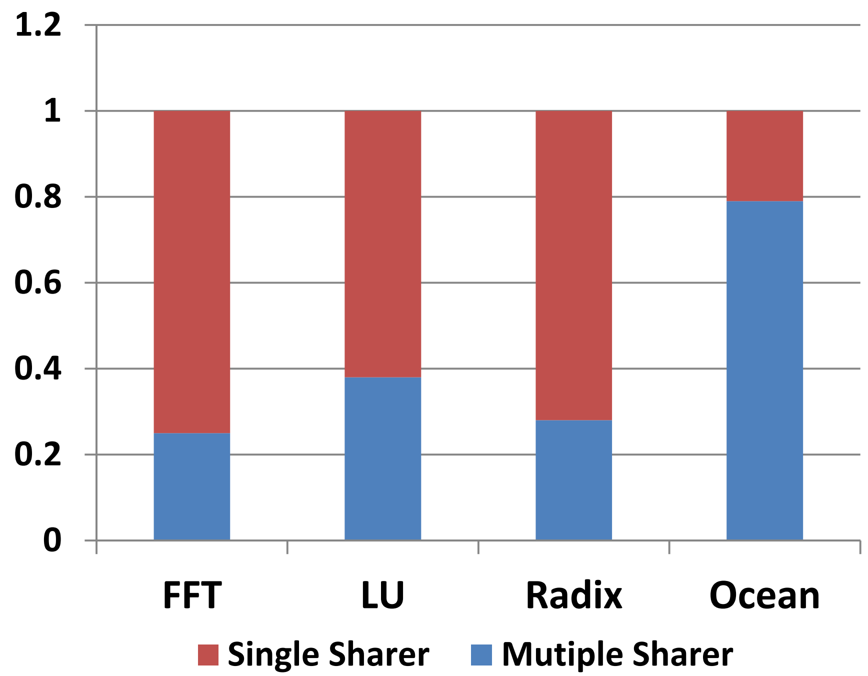
Proposed Multicast Sub-Network

For larger networks, snoopy-based cache coherence reduces performance

- Broadcasting data to all shared tiles, consuming more address bandwidth
- Consumes more latency and power at the caches

- Wavelength routed second multicast sub-network
- Filter and route cache requests to nodes that hold the cache data
- Reduction in required bandwidth and power dissipation
- Potential for simultaneous multiple requests (could lead to race conditions)

Percentage of request with multiple sharers



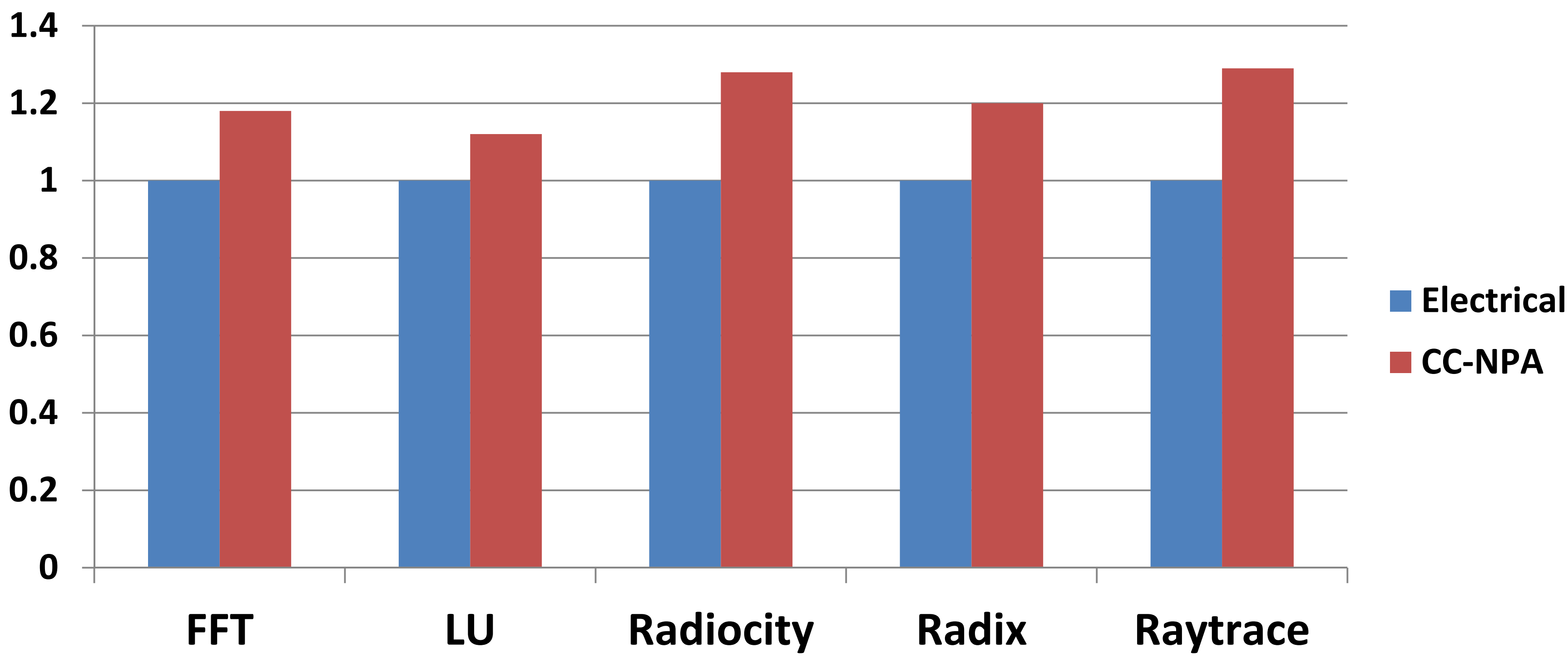
Initial Performance Analysis

- Performance Comparison
  - Simics with Gems Memory Module
  - FFT, LU, Radiosity, Ocean, Radix, & Water
- Area & Power Analysis

Simics Parameters

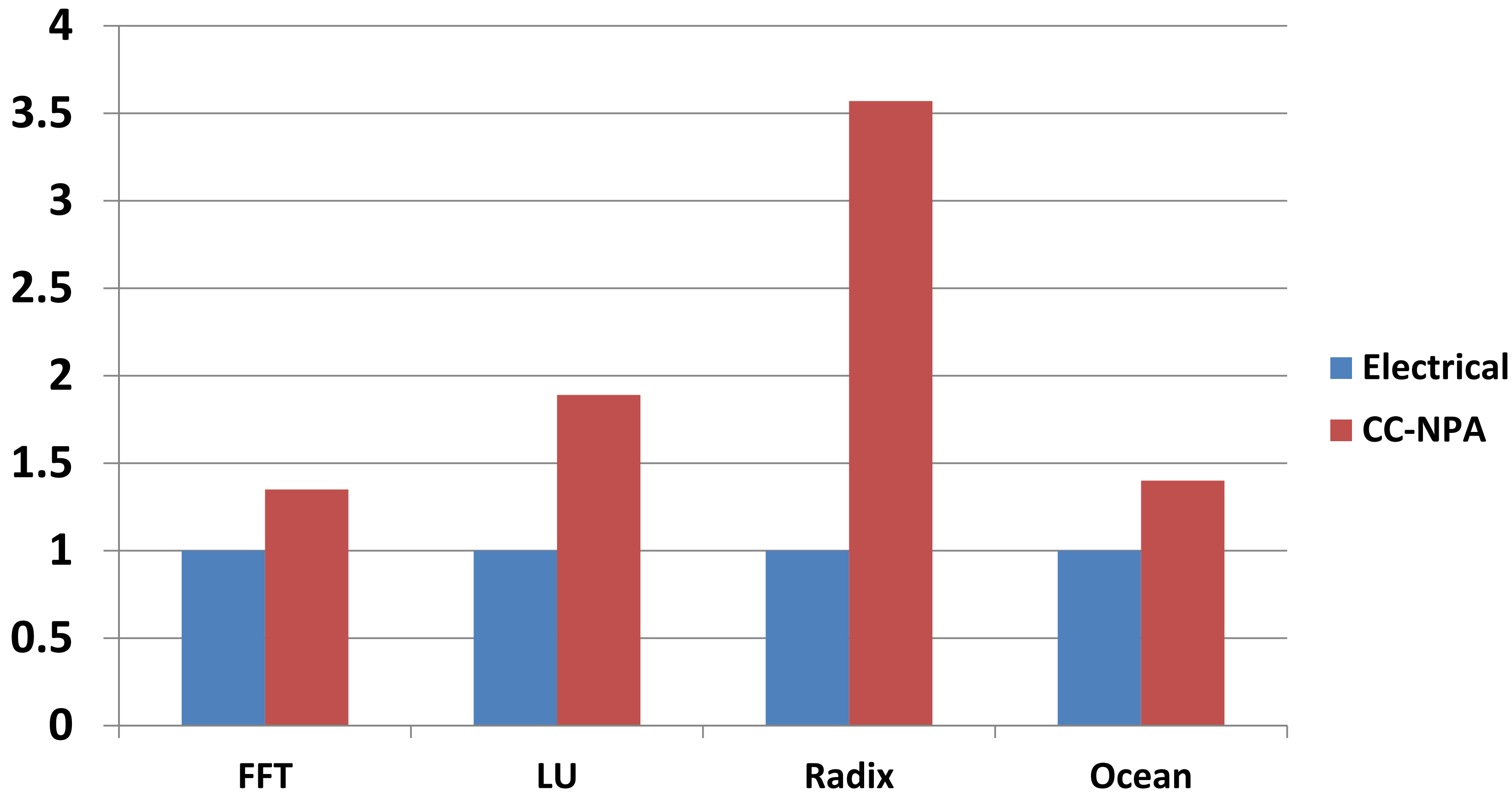
Parameter	Value	Parameter	Value
L1/L2 coherence	*MOSI*	**Core Frequency**	*5 GHz*
L2 cache size/accoc	*256 KB/16-way*	**Threads (core)**	*2*
L1 cache/accoc	*64KB/4-way*	**Issue policy**	*In-order*
Cache line size	*64B*	**Memory Size (GB)**	*4*
Memory Controllers	*16*	**Address Bandwidth (opt)**	640 GBps
Address Bandwidth (elec)	*320 GBps*		

Splash-2 Speed up (16-cores)



- CC-NPA increases performance by about 25%

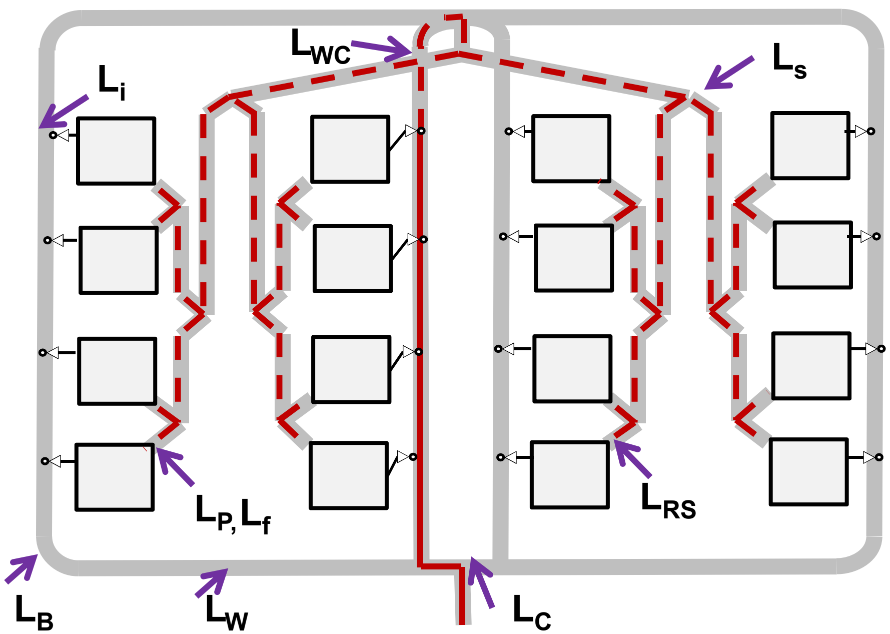
Splash-2 Speed up (64-cores)



- CC-NPA increases performance of up to 2x

Power Analysis

Device	Loss(dB)	Device	Loss(dB)
Coupler (L_c)	*1*	Filter drop (L_f)	*1*
Non-Linearity (L_n)	*1*	Bending (L_B)	*1*
Photo-detector (L_p)	*1*	Waveguide Crossing (L_{wc})	*0.05*
Modulator Insertion (L_i)	*1*	Receiver (L_{RS}) Sensitivity	*-20 dBm*
Waveguide (per cm) (L_W)	*1.3*	Splitter (L_s)	*3*
Laser Efficiently	*30%*	Ring modulation	*150 fJ/b*
Ring Heating	*100 fJ/b*	TIA/ voltage amp.	*1.1 pJ/b – 100 fJ/bit*

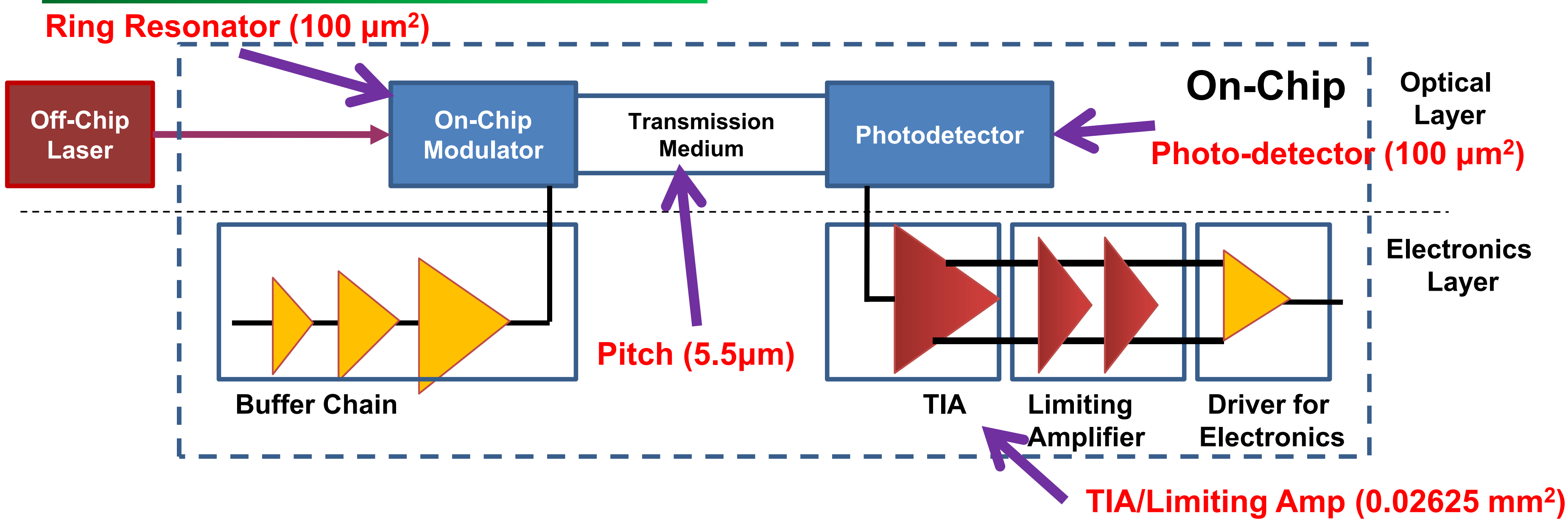


$5\times L_S + 7\times L_W + L_C + L_N + 3\times L_I + L_F + 8\times L_B + 100\times L_{WC}$

-43.1 dB (per wavelength)

Total Power (opt) = 5.44 W (8 wavelengths)

Area Analysis



Device	Area (μm^2)
Waveguide (pitch)	5.5 µm
Micro-ring resonator	100
Photo-detector	100
TIA/ Limiting Amplifier	0.02625 (mm²)

Broadcast Sub-Network: 24 mm² (optical)
51 mm² (electrical)

Conclusion & Future Work

- CC-NPA is both a low power & high bandwidth network for future cache coherent many-core processors
- CC-NPA combines the benefits the of snoopy cache coherent protocols and nanophotonics
- CC-NPA provides scalable bandwidth using two sub-networks (broadcast and multicast)
- Future work will involve designing and optimizing the multicast sub-network