McCormick
Northwestern Engineering

Towards Energy-Proportional Optical Interconnects

Nikos Hardavellas, Northwestern University
Yigit Demir, Computational Lithography, Intel

OPTICS Workshop
March 18th, 2016

partially supported by NSF award CCF-1453853

McCormick
Northwestern Engineering

Photonics Need High Power Lasers

- Emergence of photonics
  - High bandwidth, low latency, energy efficient
  - Wide range of apps: manycores, multi-chip, datacenters
- However, lasers are really power-hungry
  - Optical devices induce optical loss (13+ dB is typical)
  - WDM-compatible lasers are 5–30% efficient

➡ 10–20x higher power than required optical output

Most of the Laser Power is Wasted

Demir & Hardavellas [HPCA'15] [NOCS'15] [SPIE'15] [IPC'14] [ISLPED'14]

- Interconnect may stay idle for long times
  - Compute-intensive execution phases of workloads
  - 30% server utilization in data centers [Barroso 2007]
- But laser stays always on!
  - …even during periods of interconnect inactivity

➡ Up to 94% laser energy waste in real-world workloads

Proposed Solution: Laser Power-Gating

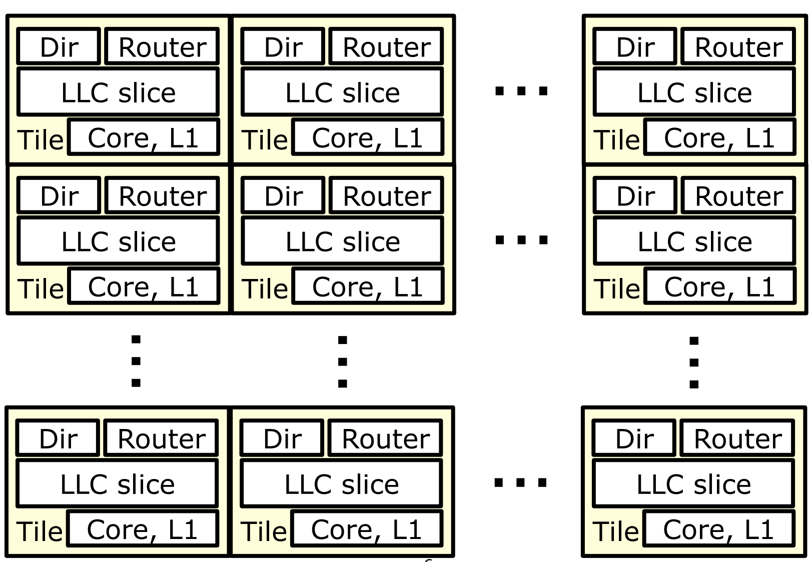
- Turn the lasers off when interconnect is idle
- Turn the lasers on before sender transmits
  - This may be tricky… needs early warning or predictive schemes
- Overlooked until recently
  - Traditional comb lasers are slow to turn on
- New enabling technology: *Fast on-off switching on-chip lasers*
  - InP, Ge, … Turn on/off in 1.5–2 ns
  - On-chip → simplify design and lower cost
  - [HPCA'15] [SPIE'15] [IPC'14] [ISLPED'14]

ProLaser: Energy-Proportional Photonic Nets.

- Power saving mechanism for photonic interconnects
  - Laser power-gating
    - → Independent power gating for data and control bits
    - → Predicts laser turn-on
    - → Saved power can be used by the cores

- Result highlights
  - Laser energy reduction: 42–88% (61% on avg.)
  - Processor energy reduction: 35–52% (40% on avg.)
  - Leads to 50–73% speedup (60% on avg.)
  - Within 2–6% of the theoretically maximum savings

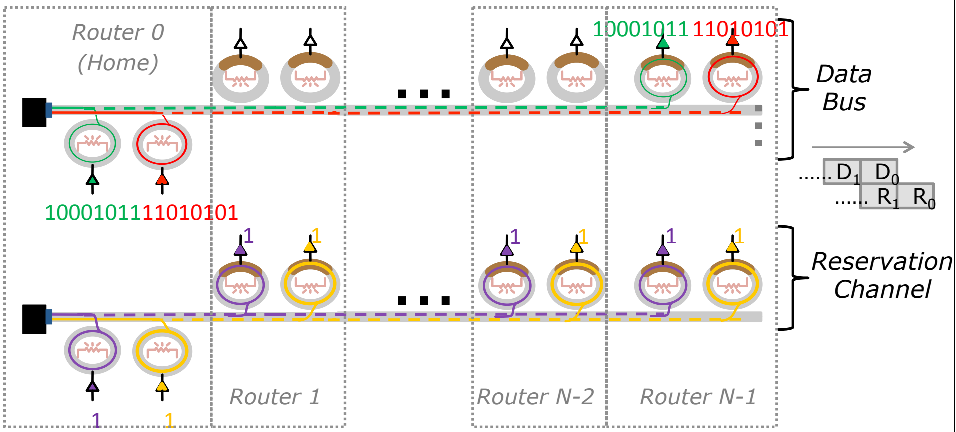
Architecture Assumed: Tiled Multicore



McCormick

Northwestern Engineering

Architecture Assumed: R-SWMR Optical Crossbar



McCormick

Northwestern Engineering

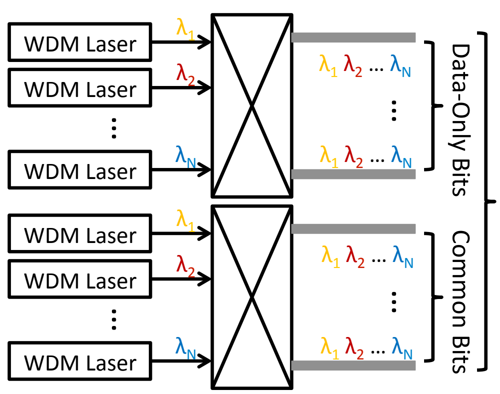
Segregating Data from Control Bits

[LaC, Demir & Hardavellas, IPC'14]

Most on-chip messages are short. On average:

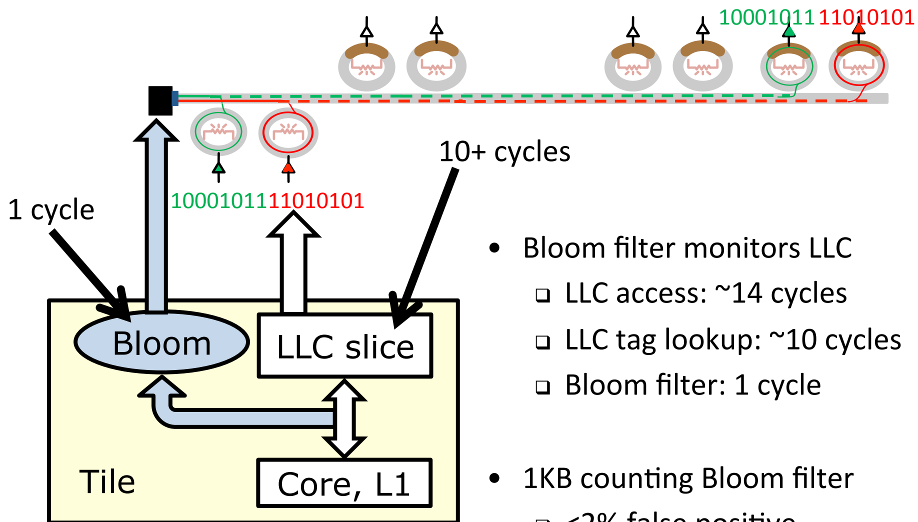
65% are Control Messages
88 bits

35% are Data Messages
600 bits



- Turn on the common bits (44) for all types of data
- Turn on the data-only bits (300) for data messages only

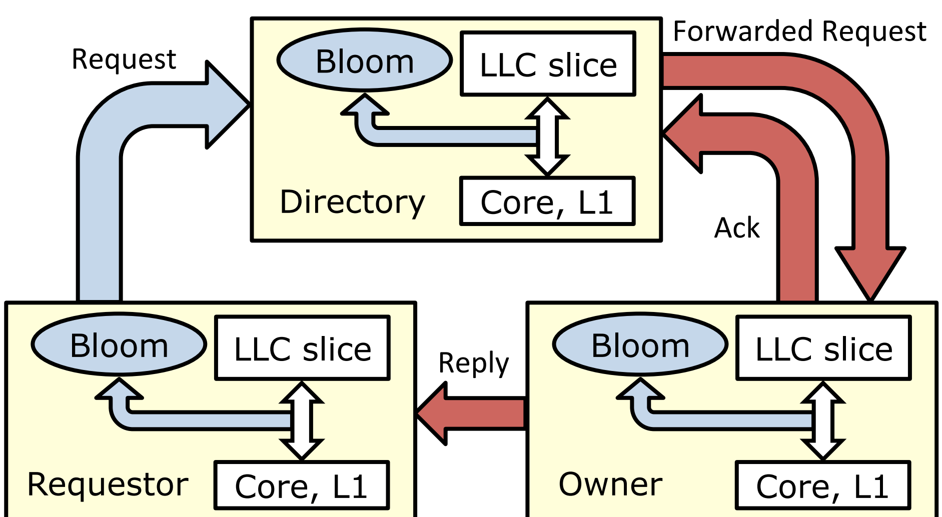
Predictive Laser Turn-On: Communication Initiation



- Bloom filter monitors LLC
  - LLC access: ~14 cycles
  - LLC tag lookup: ~10 cycles
  - Bloom filter: 1 cycle
- 1KB counting Bloom filter
  - <2% false positive

Laser Control Co-design with Coherence Protocol

[EcoLaser+, Demir & Hardavellas, SPIE'15]

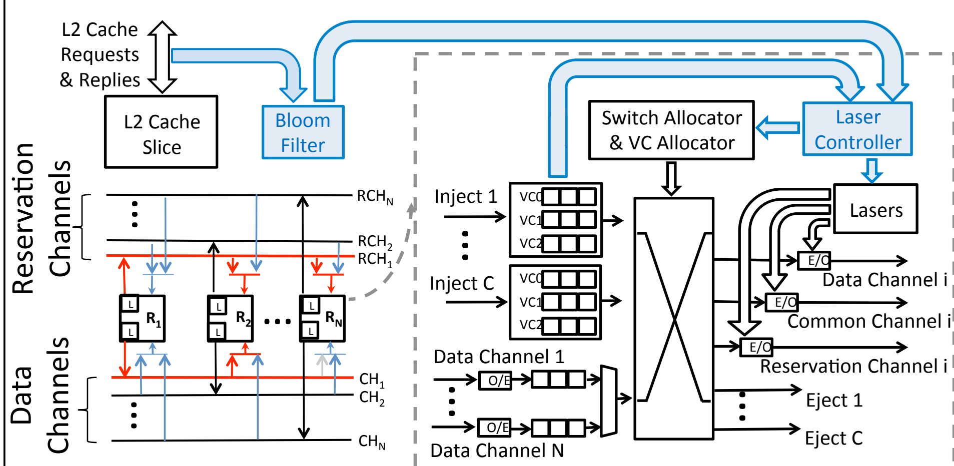


Laser Control Co-design with Coherence Protocol

[EcoLaser+, Demir & Hardavellas, SPIE'15]

- Anticipates laser activation
  - Correlates cache coherence requests to replies
  - Activates laser early → hides laser turn-on delay
- Which laser / plane to turn on?
  - Predict cache miss → turn on requestor's control plane
  - Request to directory → turn on directory's control plane
  - Directory forwarding → turn on owner's control+data plane
  - etc… (including memory controller)
- When to turn it on?
  - Turn-on the laser just 1.5ns before the payload is ready
  - Minimum latency for each operation

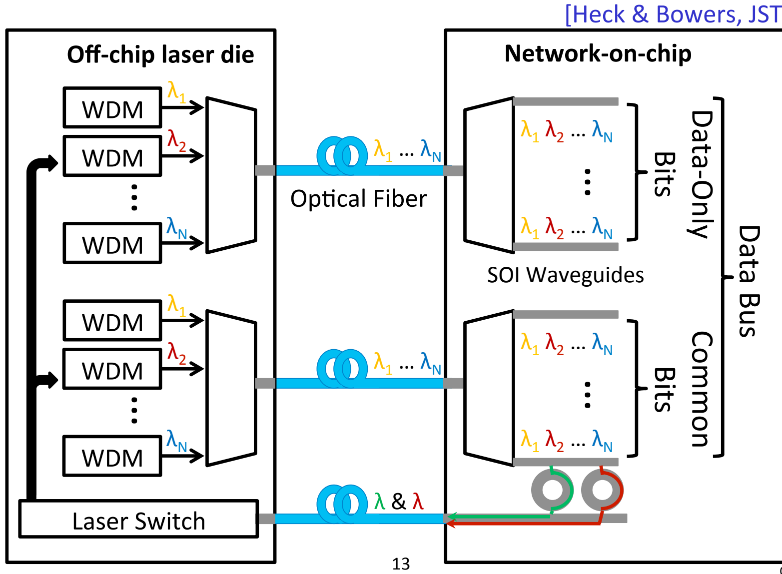
Router Microarchitecture (R-SWMR)



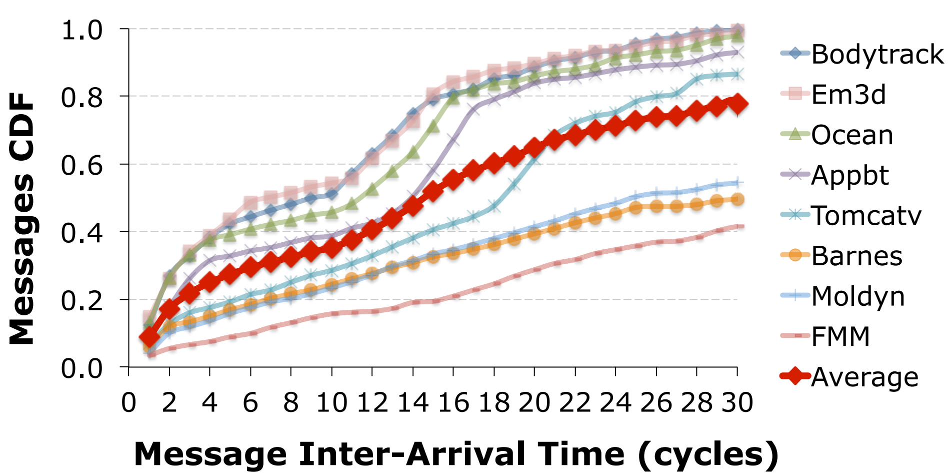
Bloom filters + coherence protocol → predict accesses

Controlling Off-chip Laser Source

[Heck & Bowers, JSTQE'14]

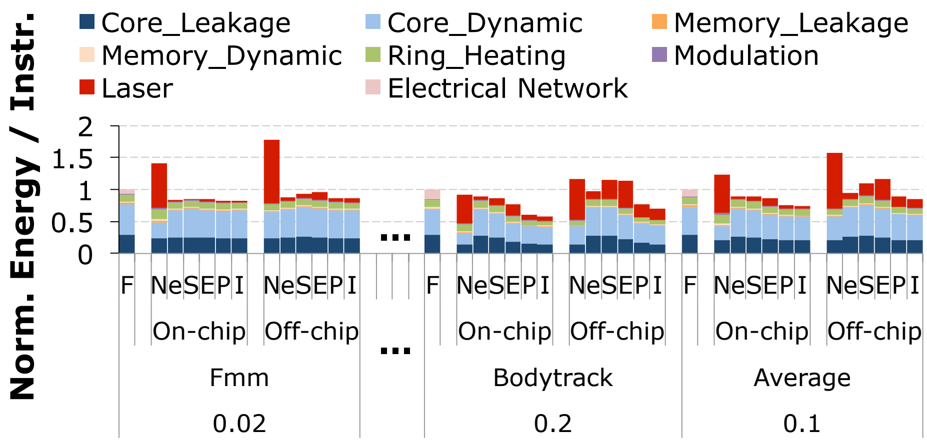


Interconnect Often Stays Idle



➡ 70% messages on avg. ≥ 9 cycles apart (1.8ns), 40% are ≥ 4ns

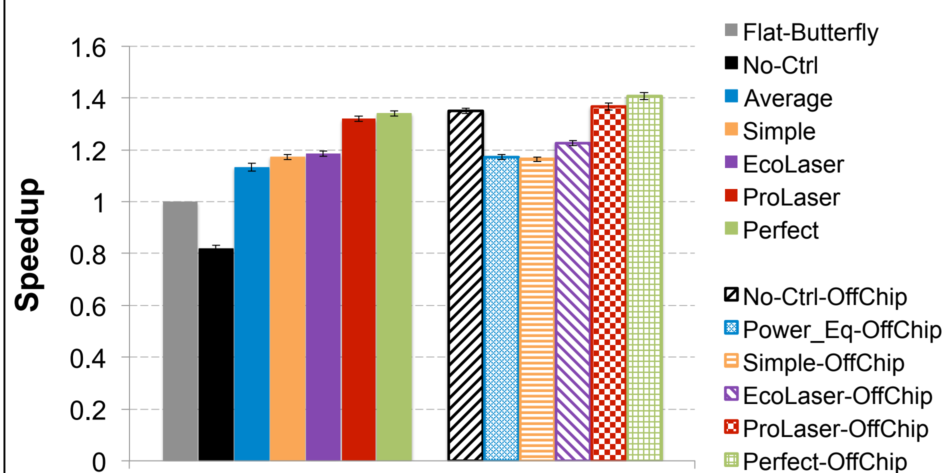
ProLaser Energy Savings



- ProLaser saves 49–88% of laser power
- 35–52% lower energy / instruction (40% avg.)

McCormick
Northwestern Engineering

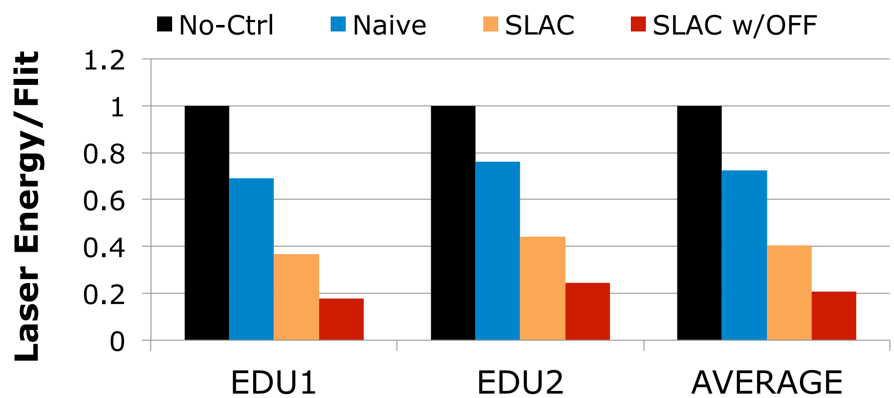
ProLaser: Performance Impact on Real Workloads



- 60% speedup over No-Ctrl; 40% over flattened buttefly

McCormick
Northwestern Engineering

Teaser Slide: Laser Gating in the Datacenter



- SLaC w/OFF optimization saves 79% of the laser energy
- Similar results with Facebook and Microsoft traces

McCormick
Northwestern Engineering

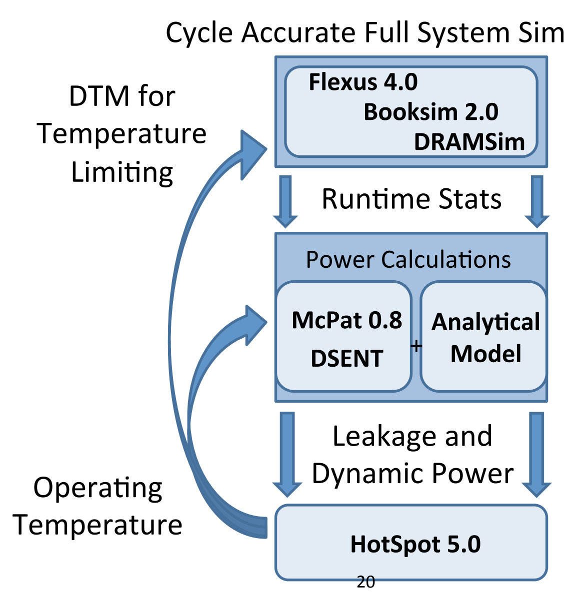
Conclusion

- Problem: lasers are really power hungry, mostly wasted power
- Our solution: laser power-gating (**ProLaser**, **SLaC**, EcoLaser (+), LaC)
- Significant energy reduction
  - Laser: 42–88% (61% on avg.), Processor: 35–52% (40% on avg.)
  - Within 2–6% of the theoretically maximum savings
  - Power reduction leads to speedups: 50–73% (60% on avg.)
- Applicable to a wide range of scales (on chip, multichip, datacenter)

Thank you! **Questions?**

Backup Slides

Simulation Tool Chain



Experimental Methodology

CMP Size	64 cores, 580 mm^2
Core	ULTRASPARC III ISA, up to 5Ghz, OoO, 4-wide dispatch/retirement, 96-entry ROB
L1 Cache	Split I/D, 64KB 2-way, 2-cycle load-to-use, 2 ports, 64-byte blocks, 32 MSHRs, 16-entry victim cache
L2 Cache	512 KB per core, 16 way, 64-byte blocks, 14 cycle-hit, 32 MSHRs, 16-entry victim cache
Memory Controller	One per 4 cores, 1 channel per Memory Controller, Round-robin page interleaving
Main Memory	Optically connected memory [3], 10ns access
Network	R-SWMR radix-16 crossbar and firefly, 300-bit wide links @ 10GHz, 20 flit deep buffers, 3 cycle router delay

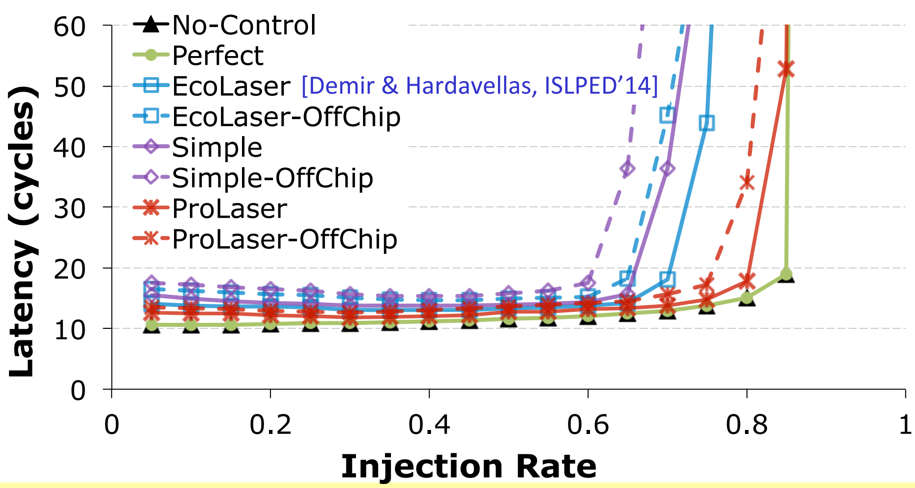
Nanophotonic Parameters

Radix-16 SWMR		**On-Chip Laser**	**Off-Chip Laser**
	per Unit	**Total**	**Total**
DWDM		64	64
Splitter	0.2 *dB*	0.6 *dB*	0.6 *dB*
WG Loss	0.3 *dB/cm*	3 *dB*	3 *dB*
Nonlinearity	1 *dB*	1 *dB*	1 *dB*
Modulator Ins.	0.5 *dB*	0.5 *dB*	0.5 *dB*
Ring Through	0.01 *dB*	10.24 *dB*	10.24 *dB*
Filter Drop	1.2 *dB*	1.2 *dB*	1.2 *dB*
Coupler	2 *dB*		4 *dB*
Total Loss		**16.64 *dB***	**20.64 *dB***
Detector		**-20 *dBm***	**-20 *dBm***
Laser Power Per Wavelength		**0.461 *mW***	**1.158 *mW***
Total LaserPower	*15% Eff.*	**14.8 *W***	**37.1*W***

Workloads

- Fmm: Input 128K
- Moldyn: 15, 20, 3.2 M
- Barnes: Input 64K
- Tomcatv: 4096, 10
- Appbt: in.24x24x24x8bit
- Ocean: 1026, 9600
- Em3d: 400K, 2, 15, 5
- Bodytrack

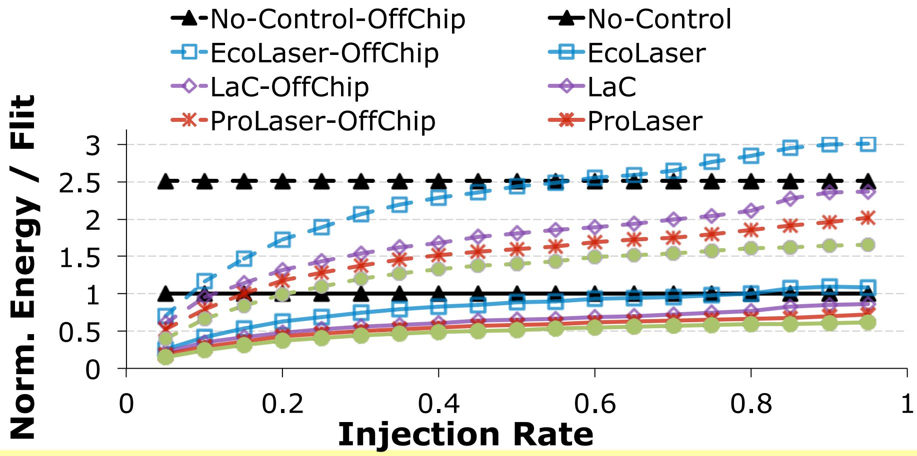
Interconnect Performance



➡ ProLaser almost perfect saturation; EcoLaser saturates early

McCormick
Northwestern Engineering

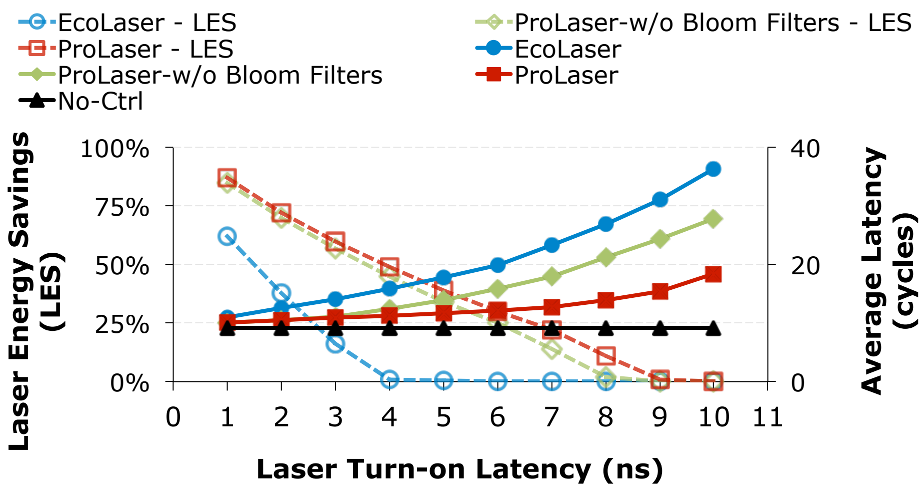
Interconnect Energy



- ProLaser saves 49–88% of laser power
- ProLaser is ~2x better than EcoLaser; 2–6% of Perfect

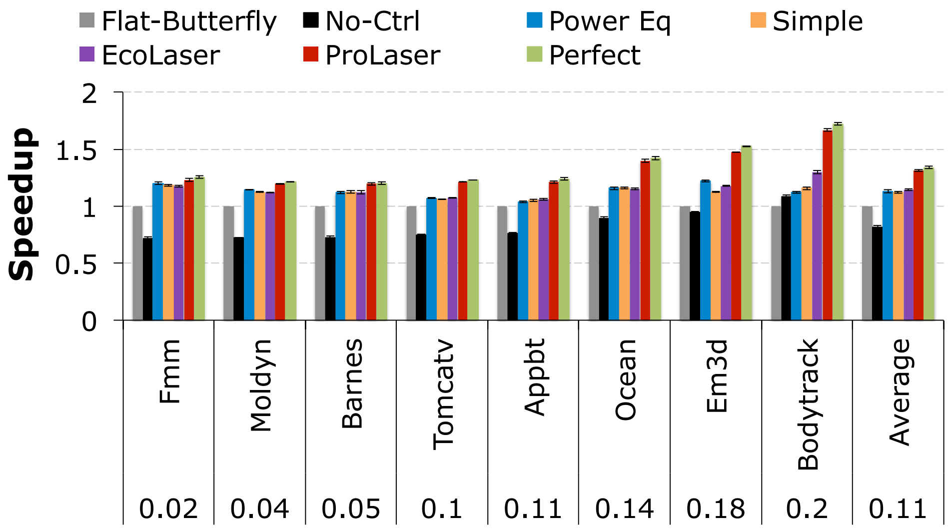
McCormick
Northwestern Engineering

Sensitivity to Laser Turn-On Delay

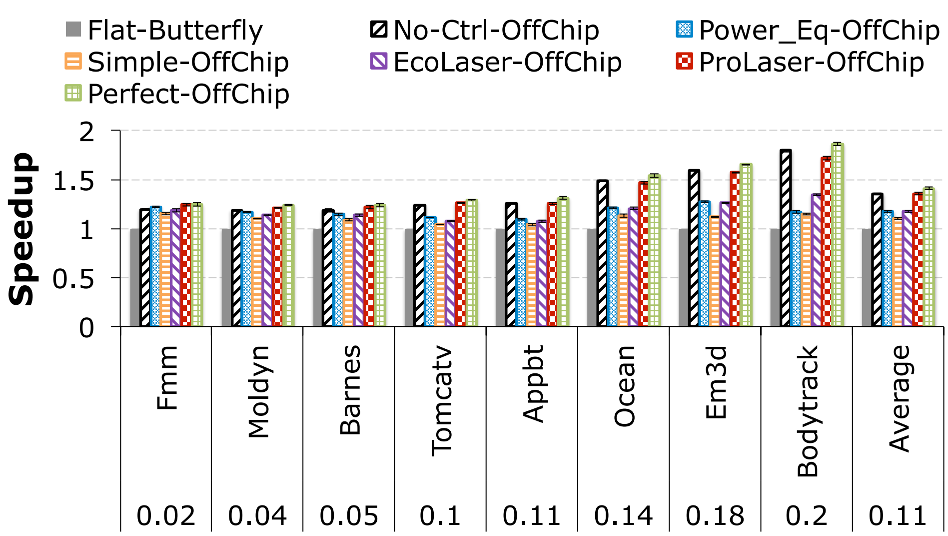


- ProLaser tolerates 2.3x higher laser turn-on delay than EcoLaser

Speedup (On-chip Lasers)

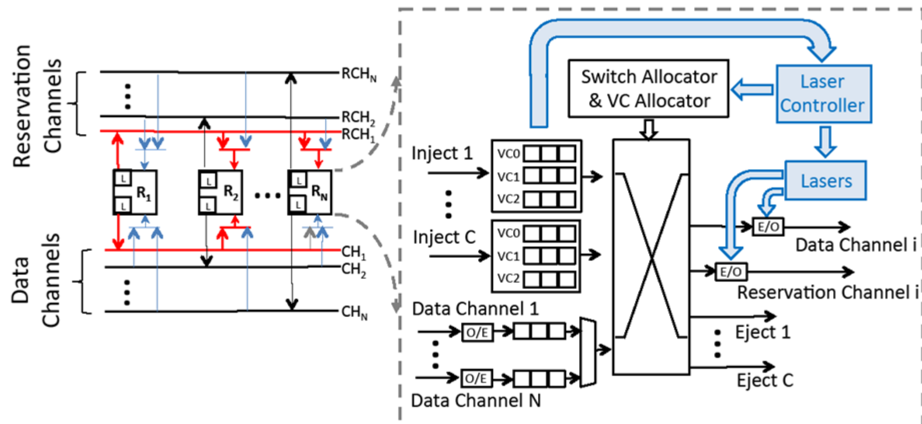


Speedup (Off-chip Lasers)



Router microarchitecture for SWMR EcoLaser

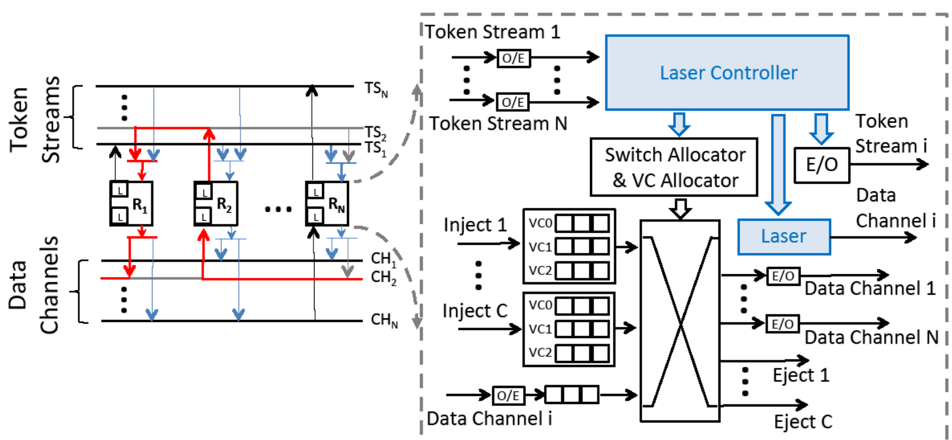
[Ecolaser, Demir & Hardavellas, ISLPED'14]



Message in injection buffers → Laser Turn On

Router Microarchitecture for MWSR EcoLaser

[Ecolaser, Demir & Hardavellas, ISLPED'14]



Adaptive Laser Turn-on + Specialized Token Stream

MWSR Optical Bus