Thermal Management of Manycore Systems with Photonic NoCs

Ayse K. Coskun

Boston University, ECE Department

In collaboration with:

Ajay Joshi[1], Andrew B. Kahng[2,3], Jonathan Klamkin[4],

Tiansheng Zhang[1], John Recchio[2], Vaishnav Srinivas[2], Anjun Gu[3], Yenai Ma[1]

[1]Boston University ECE Dept.;

UCSD [2]ECE and [3]CSE Dept.;

[4]UCSB ECE Dept.

This research has been partially funded by the NSF grants CNS-1149703 and CCF-1149549.
Work at UCSD has been supported by NSF, Samsung and the IMPACT+ Center.

Silicon Photonics – Benefits & Challenges

- **Silicon Photonic Links vs. Electrical Links**

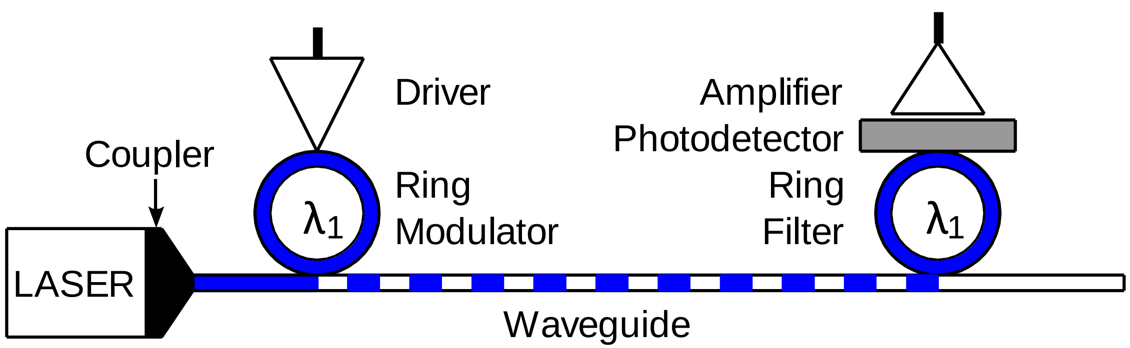
- Higher bandwidth density
- Lower long-distance communication latency
- Lower data-dependent energy consumption

- More sensitive to thermal variations
- More sensitive to process variations

High thermal tuning power

- High laser source power consumption
- Hard to package off-chip laser sources

- **Silicon Photonic Link**



Silicon Photonics – Benefits & Challenges

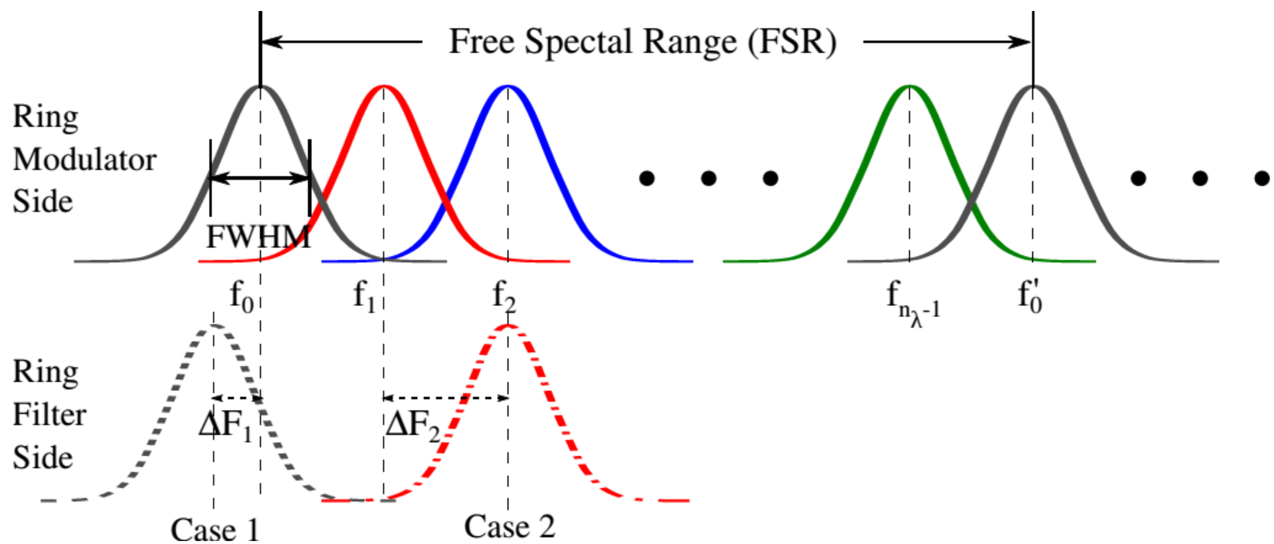
- **Thermal and Process Sensitivity of Photonic Devices**

Thermal variations → (Thermo-optic effect) → Shift in optical frequency

Process variations → Ring dimensions

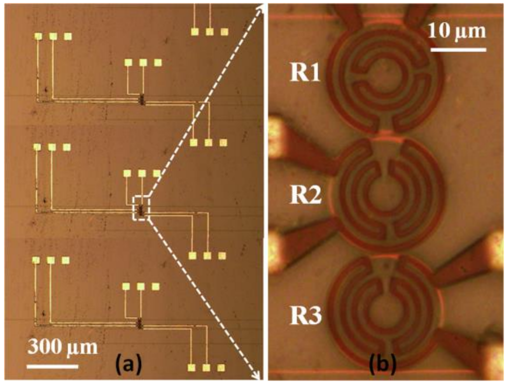
→ Optical Frequency Mismatch

- **Impacts of Optical Frequency Mismatch**



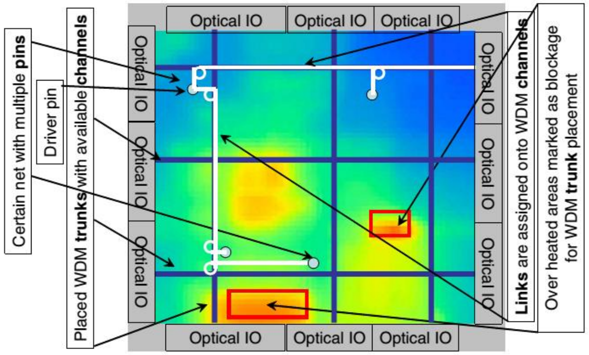
Previous Work

- Using hardware for ring filters/modulators thermal control
  - Cladding [*Djordjevic, Optical Exp.' 2013*]
  - Heaters [*Zhou, TACO'2010; Li, TVLSI'2012*]
  - Mach-Zehnder interferometers [*Biswajeet, Optical Exp.'2010*]



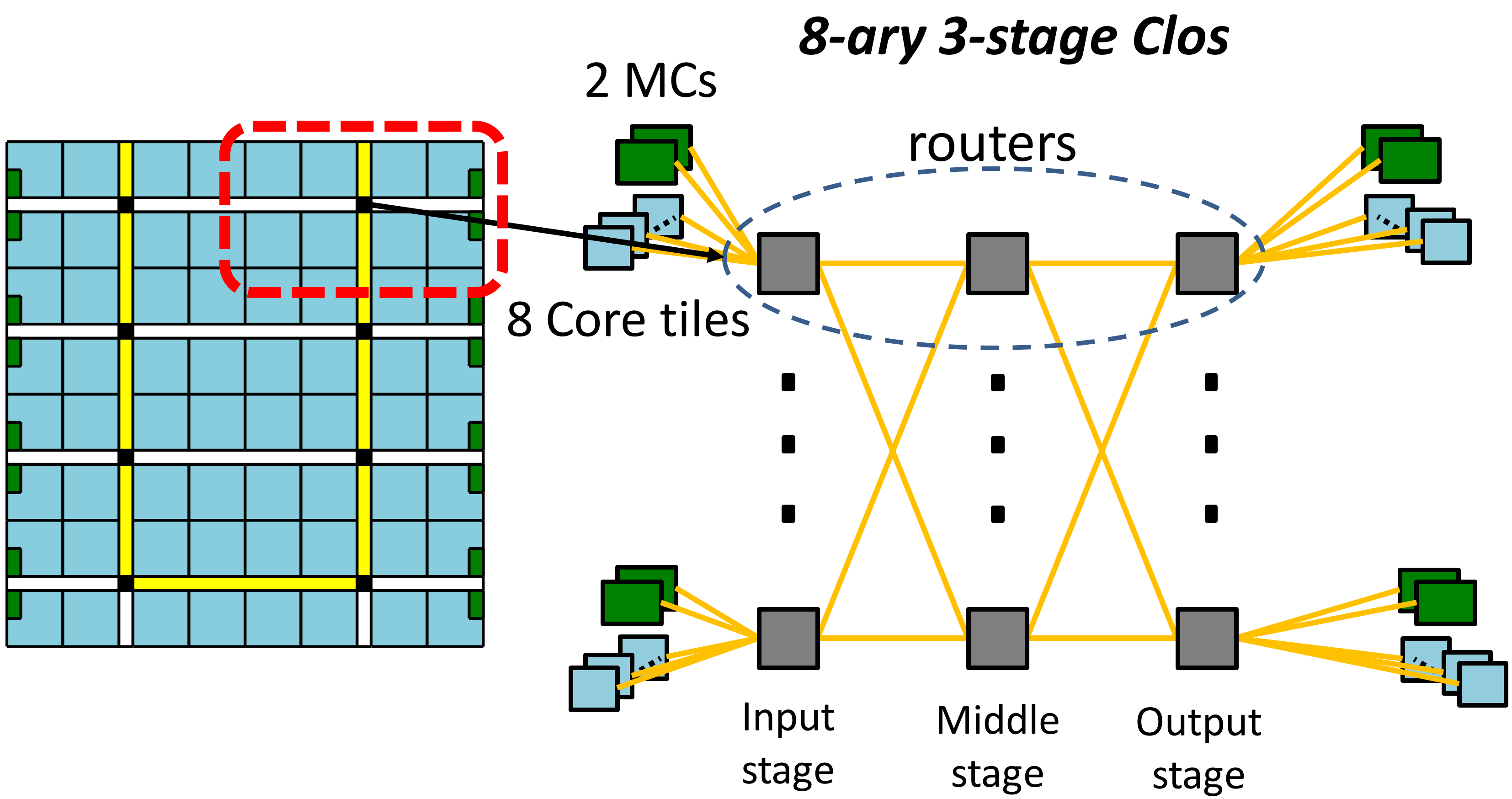
- Runtime management in manycore systems w/ PNoC
  - Aurora [*Li, TCAD'2015*]: Thermal Tuning + DVFS + Routing Algorithm

- Place & route for PNoCs
  - PROTON: An automatic tool for PNoC P & R *[Boos, 2013]*
  - GLOW: A ILP based global router for PNoC *[Ding, 2012]*

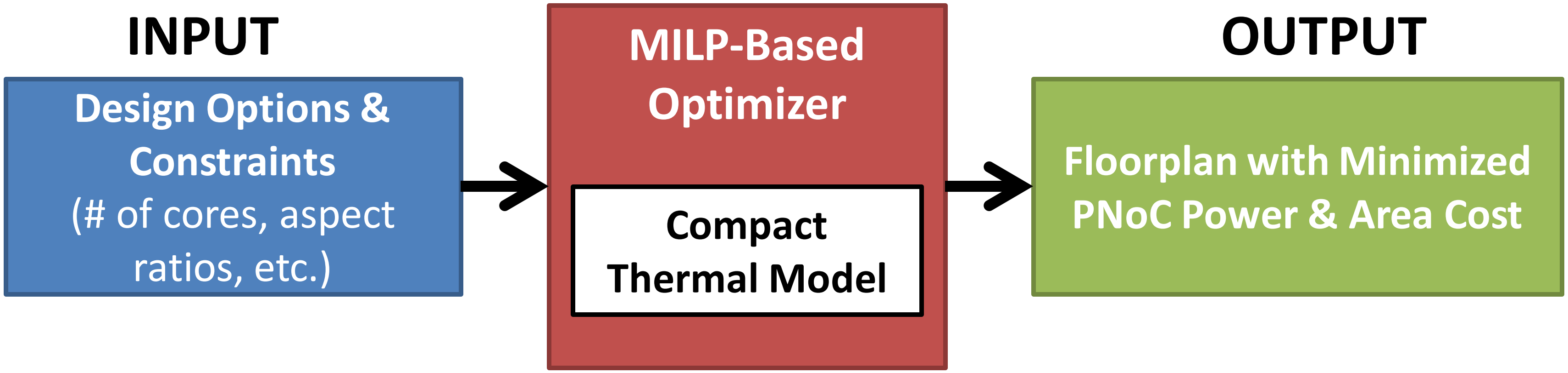


Target Manycore System w/ PNoC

- 256-core system with Clos network



Floorplan Optimization Flow



- **Optimization Goal:**

$$\textbf{Minimize:} \quad \alpha \cdot P_{PNoC} + \beta \cdot AREA_{PNoC}$$

$$P_{PNoC} = P_{laser} + P_{tuning} + P_{electrical}$$

  - **PNoC Power:**
    - P & R's impact on waveguide length, crossing and bending
    - Laser source efficiency
    - PNoC placement's impact on thermal tuning power
  - **PNoC Area:**
    - Area cost of router groups and waveguides

[DATE'16]

Floorplan Optimization Flow

INPUT

Design Options & Constraints
(# of cores, aspect ratios, etc.)

→

MILP-Based Optimizer

Compact Thermal Model

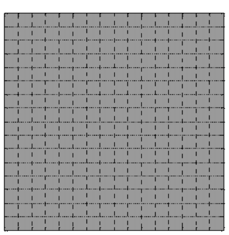
→

OUTPUT

Floorplan with Minimized PNoC Power & Area Cost

- ***Compact thermal model***

Power profile:

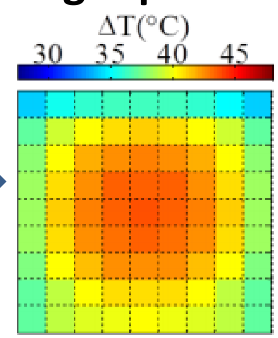


×

Compact Thermal Model

→

Accumulated thermal weight profiles



→ **Resonant frequency difference among router groups** → **Thermal tuning power**

$$[P_0, P_1, \ldots, P_{N-1}] \begin{bmatrix} \omega_{(0,0)}, & \omega_{(0,1)}, & \ldots, & \omega_{(0,M-1)} \\ \omega_{(1,0)}, & \omega_{(1,1)}, & \ldots, & \omega_{(1,M-1)} \\ \\ \omega_{(N-1,0)}, & \omega_{(N-1,1)}, & \ldots, & \omega_{(N-1,M-1)} \end{bmatrix} [\theta_0, \theta_1, \ldots, \theta_{M-1}]$$

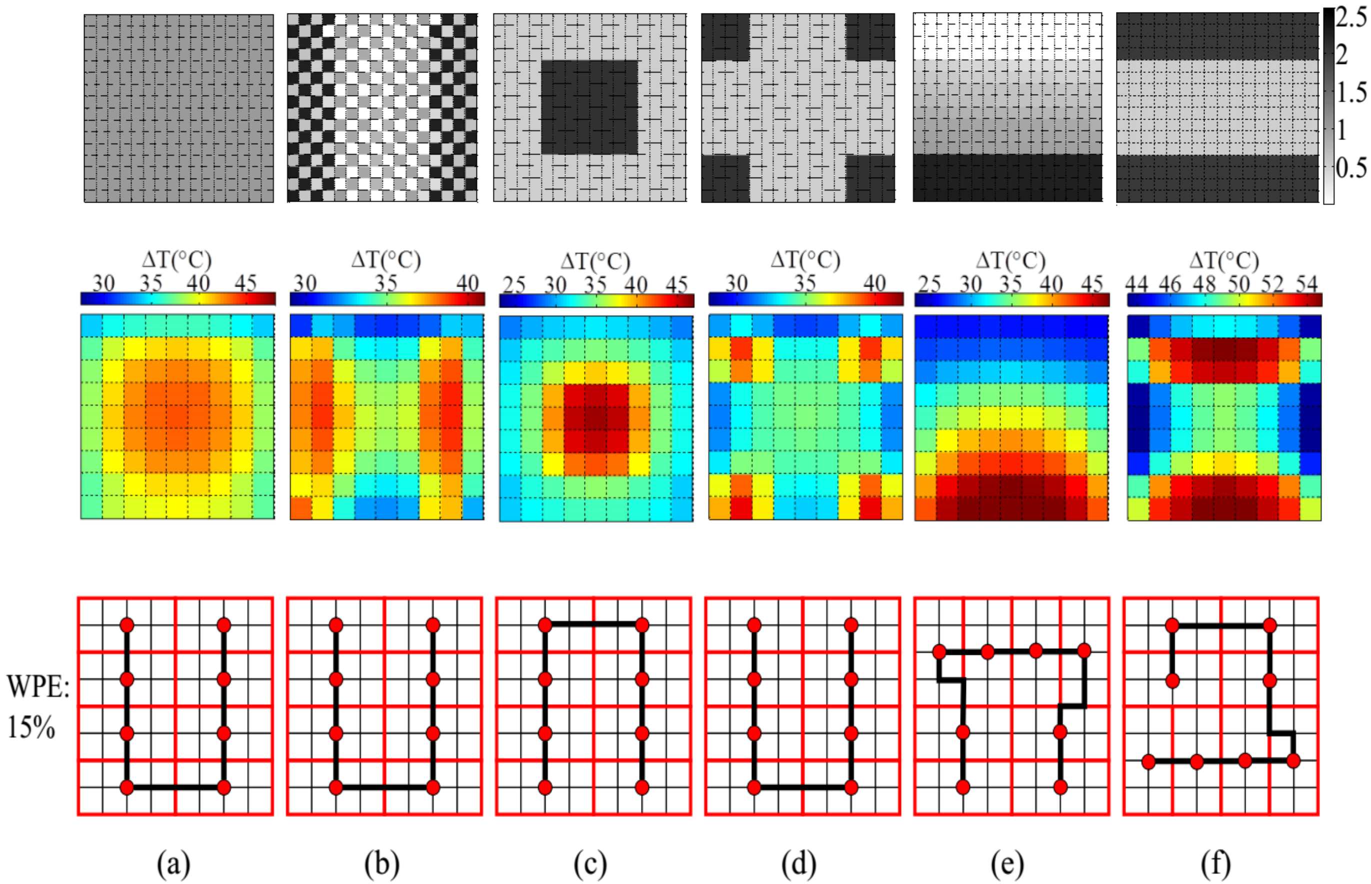
Size: $1\times N$ $N\times M$ $1\times M$

Cross-layer PNoC P&R Optimization

Power Profiles

Thermal Conditions of Potential Ring Group Locations

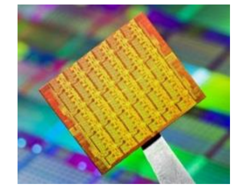
PNoC Layouts w/ Minimum PNoC Power

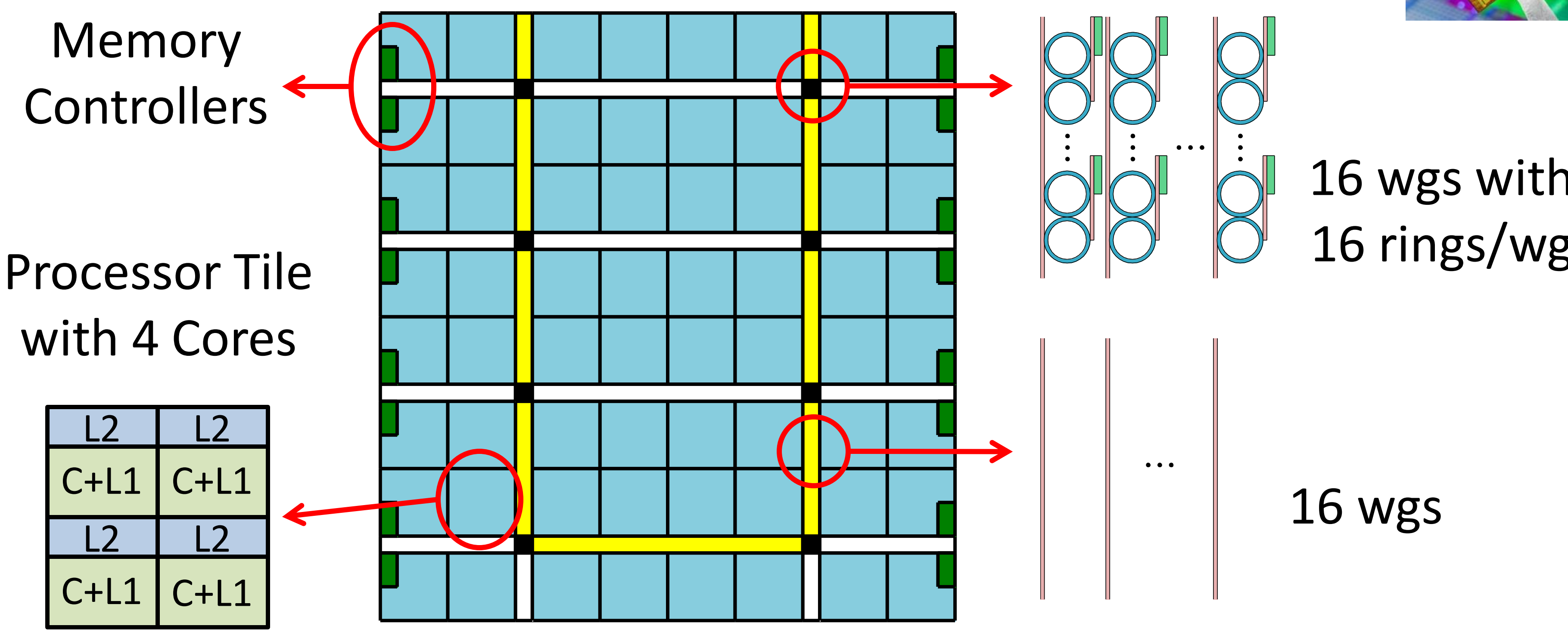


Runtime Management on Many-Core Systems with PNoC

- 256-core system with Clos network

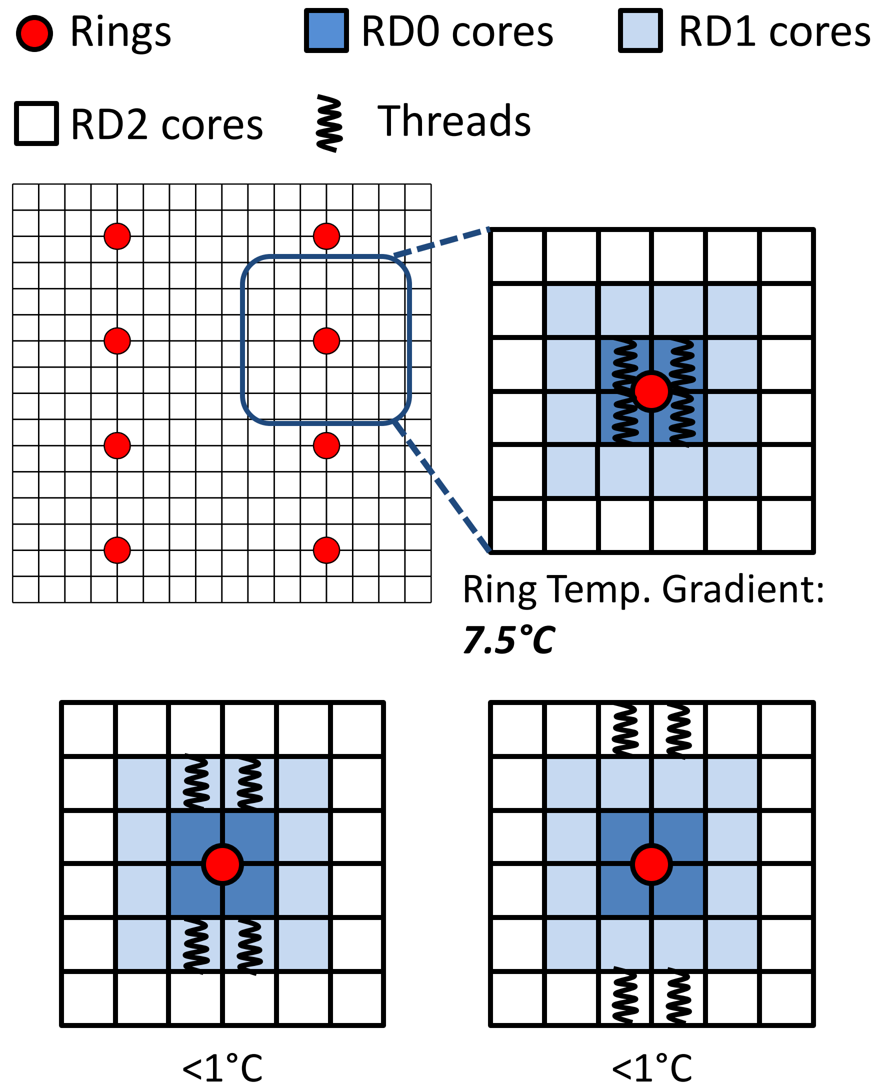
Core Architecture: IA-32 core in Intel SCC [*Howard,ISSCC2011*], 16KB I/D L1 cache & 256KB L2 cache;





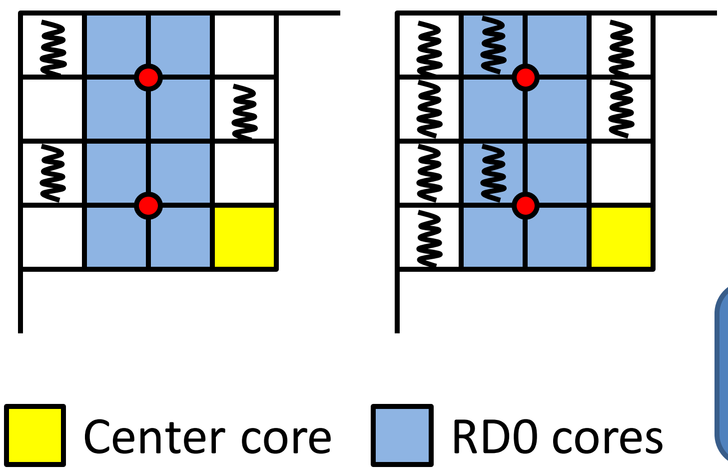
RingAware Workload Allocation Policy

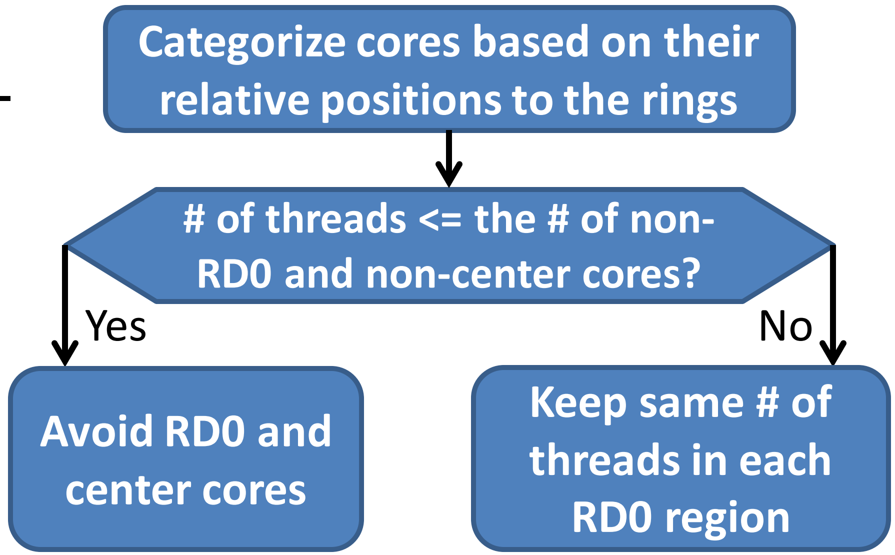
- Goals:
  - Minimize the difference among ring temperatures
  - Reduce the overall chip temperature
- Active cores' impact on ring temperature
  - Classify the cores based on their distances to a ring group



RingAware Workload Allocation Policy

- Ring temperature gradient minimization – ***RingAware***
  - Take ring locations into consideration



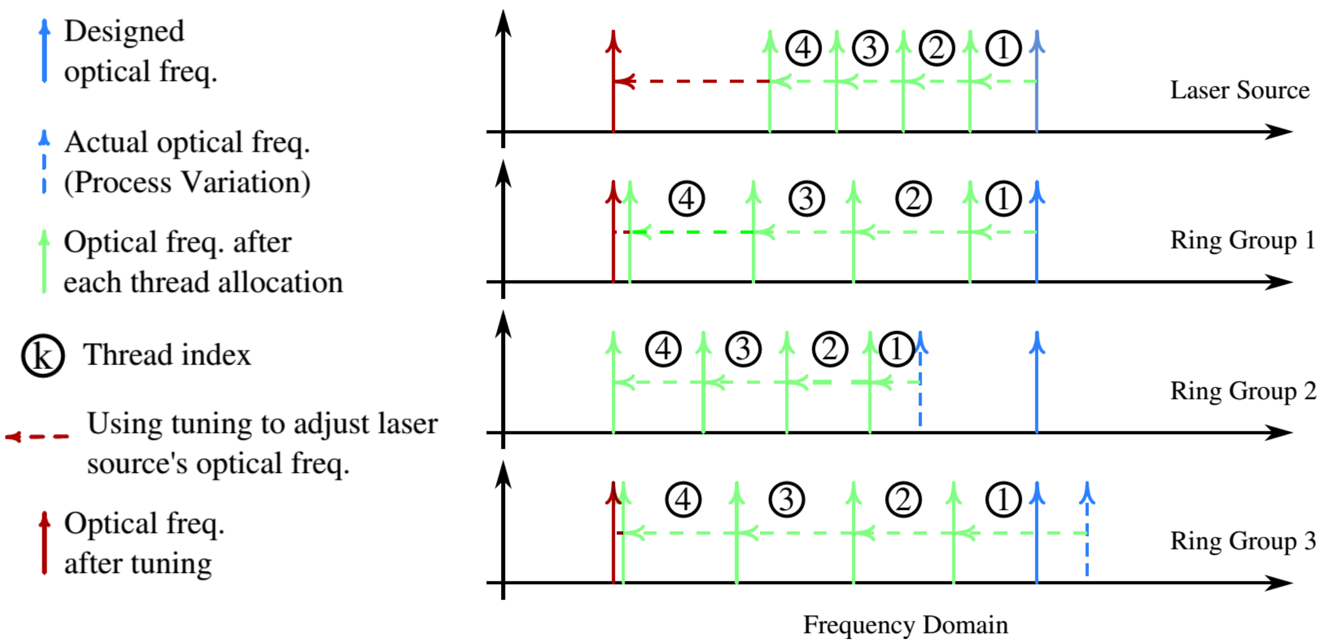


- ***Multi-program support***
  - Sort the threads based on their power dissipation & allocate high-power application first

FreqAlign Workload Allocation Policy

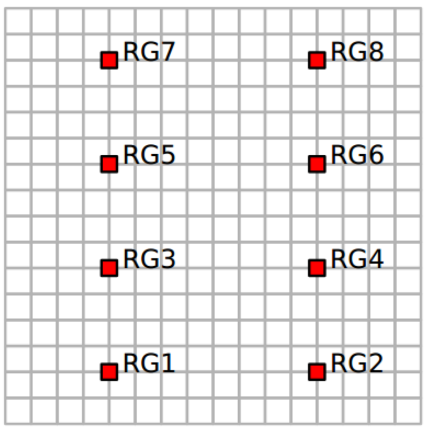
- Process variation introduces resonant frequency shift after the system is manufactured
- Only balancing the temperature of ring groups is not enough to compensate the frequency mismatch
- On-chip laser sources' optical frequencies also need to match with corresponding rings' resonant frequency

FreqAlign* + *Adaptive Frequency Tuning



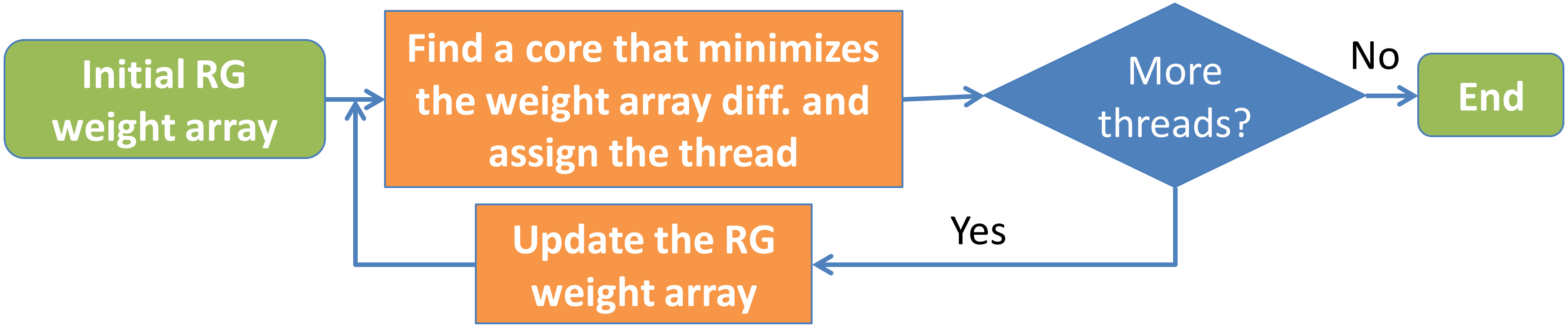
FreqAlign Workload Allocation Policy

- **Target many-core system:**



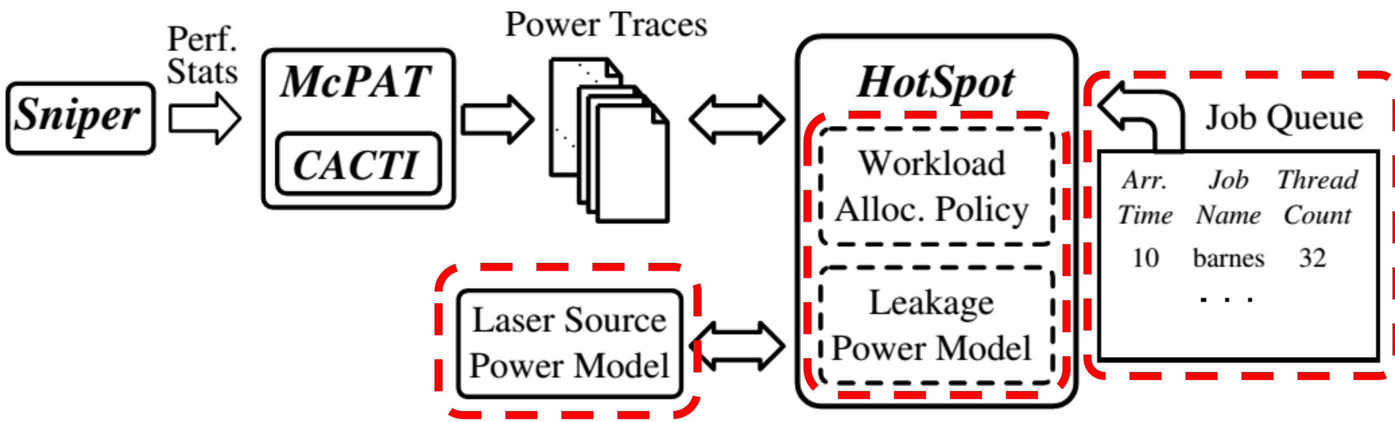
- **Two new concepts:**
  - **Weight matrix:** Every core's thermal impact on every ring group (extracted from HotSpot);
  - **Weight array:** The current frequency shift for all ring groups. Process variation is accounted when initializing this weight matrix.

- **FreqAlign Allocation Policy:**



Experimental Methodology

- **Simulation Framework:**



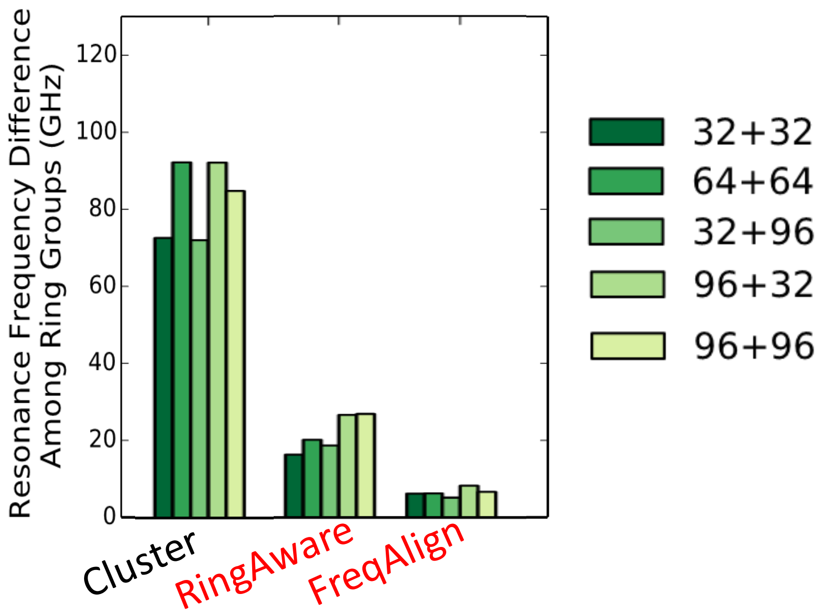
- **Workload Sets:** Selected benchmarks from SPLASH2, PARSEC and UHPC:

Workload Sets	Job 1	Job2
HP + HP	md	shock
HP + MP	md	blackscholes
HP + LP	shock	lu_cont
MP + MP	barnes	blackscholes
MP + LP	barnes	water_nsq
LP + LP	lu_cont	canneal

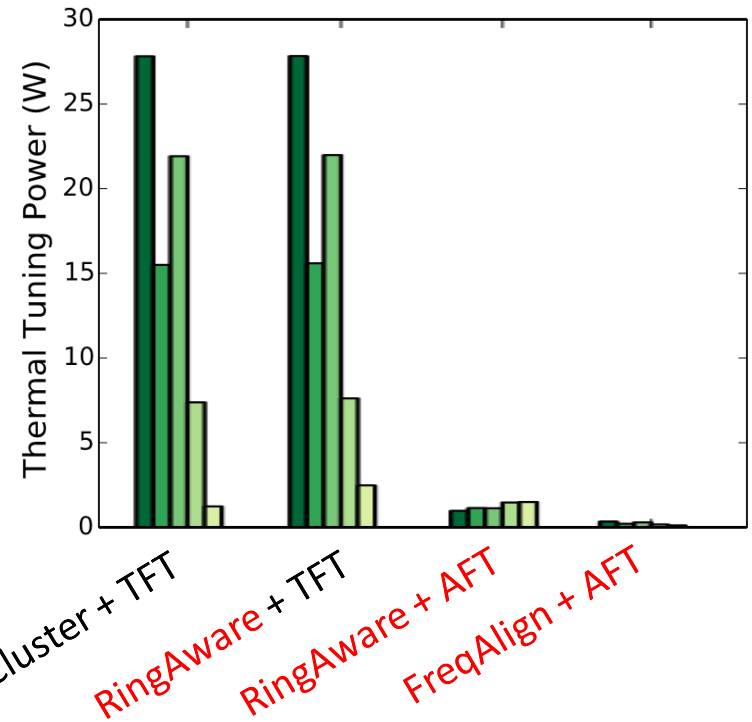
- **Tested Policies:**
  - Workload Allocation Policy: **Cluster**, **RingAware**, **FreqAlign**
  - Thermal Tuning Policy: **Target Frequency Tuning (TFT)**, **Adaptive Frequency Tuning (AFT)**

Experimental Results for Many-core System w/o Process Variations

Resonance Frequency Difference



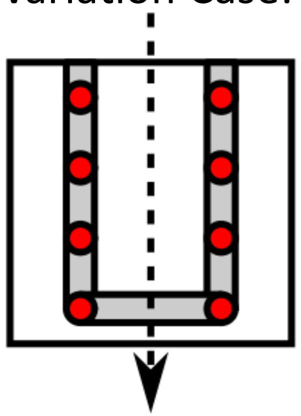
PNoC Thermal Tuning Power

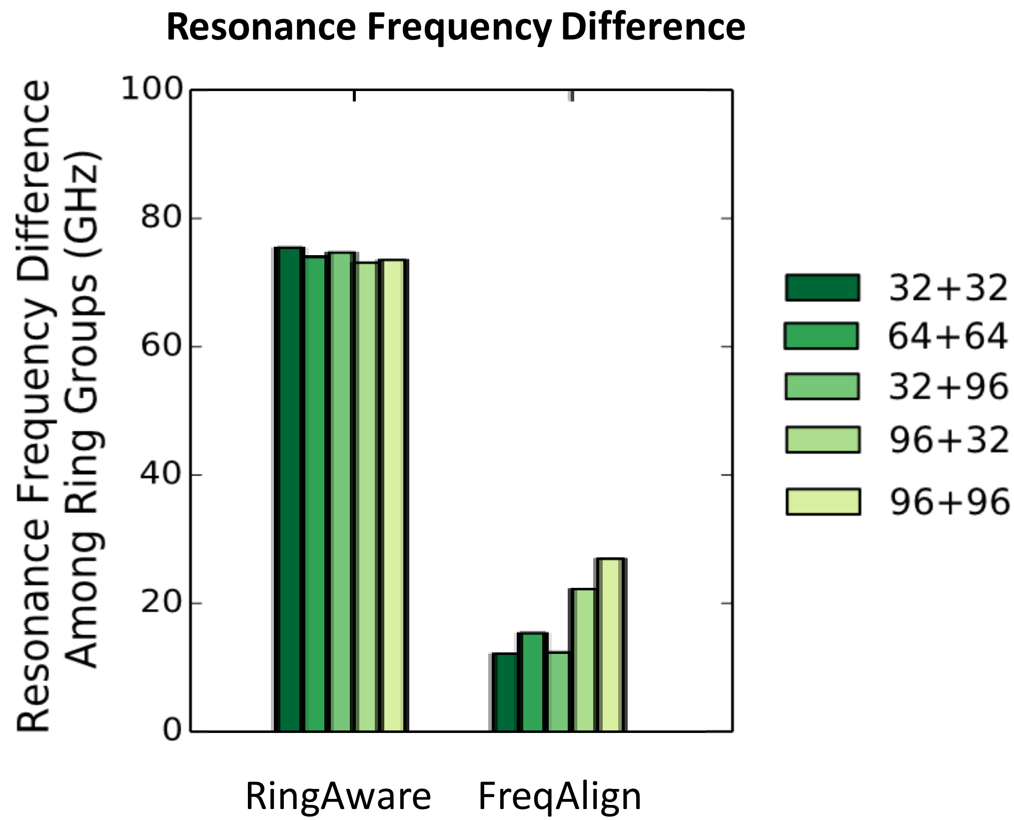


- Compared to *RingAware*, *FreqAlign* reduces the resonant frequency difference by **60.6%** on average;
- Compared to *RingAware + TFT*, *FreqAlign + AFT* reduces the tuning power by **14.93W** on average.

Experimental Results for Many-core System w/ Process Variations

Vertical Process Variation Case:





- Average resonant frequency reduction: **52.7%**
- Average tuning power reduction: **65.8%**

Laser Source Placement and Sharing

- **Laser Source Power Consumption (P_{IN})**

$$P_{IN} = \sum_{i=1}^{m} P^i_{IN} = \sum_{i=1}^{m} \frac{P^i_{OUT}}{\eta_{WPE}(T^i_{LS}, P^i_{OUT})}$$

m: # of laser sources

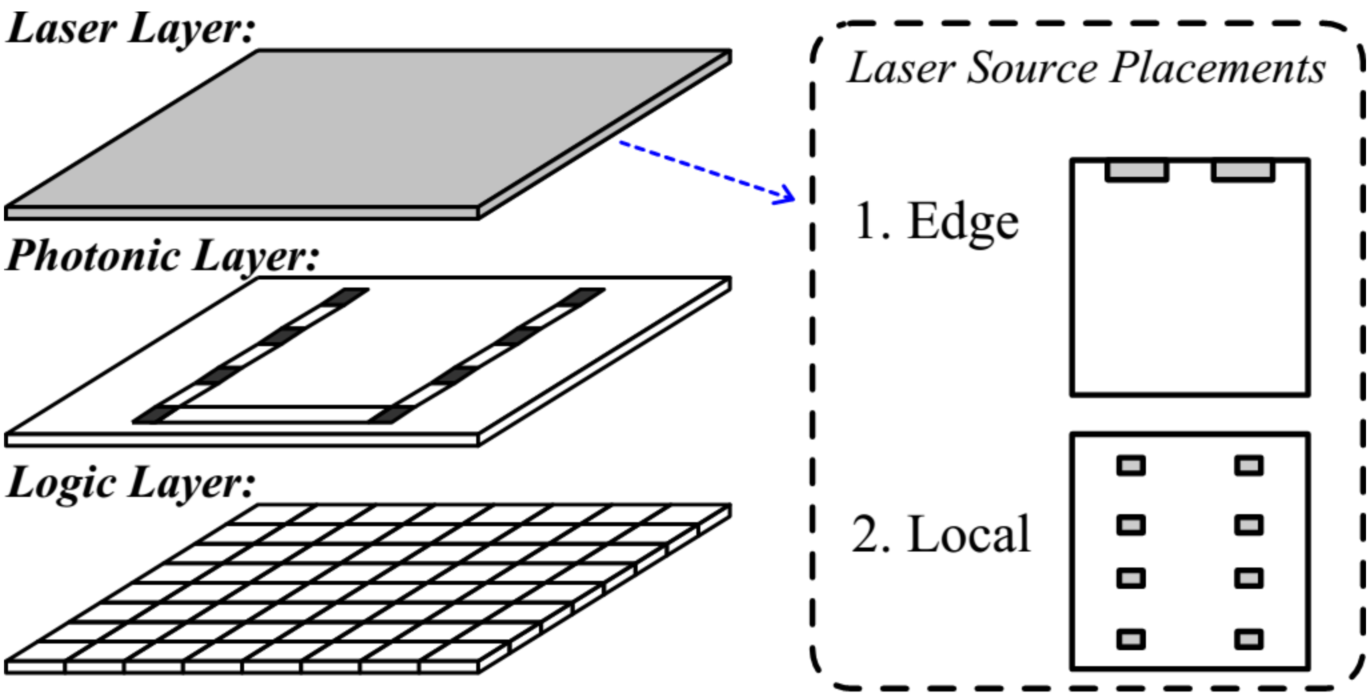
  - Wall plug efficiency (η_{WPE})
  - Operating temperature of laser sources (T_{LS})
  - Required output optical power from laser sources (P_{OUT})

$$P_{OUT} \propto Optical\ loss \begin{cases} Waveguide \begin{cases} Propagation\ loss \\ Cross/bend\ loss \end{cases} \\ Rings \\ Photodetector \end{cases}$$

- **Laser Source Power Reduction**
  - Reduce ***propagation loss*** (thus, P_{OUT}) by placing laser sources locally to rings
  - Reduce # of laser sources and increase η_{WPE} by sharing laser sources among waveguides

[NOCS'14]

Laser Source Placement and Sharing Examples

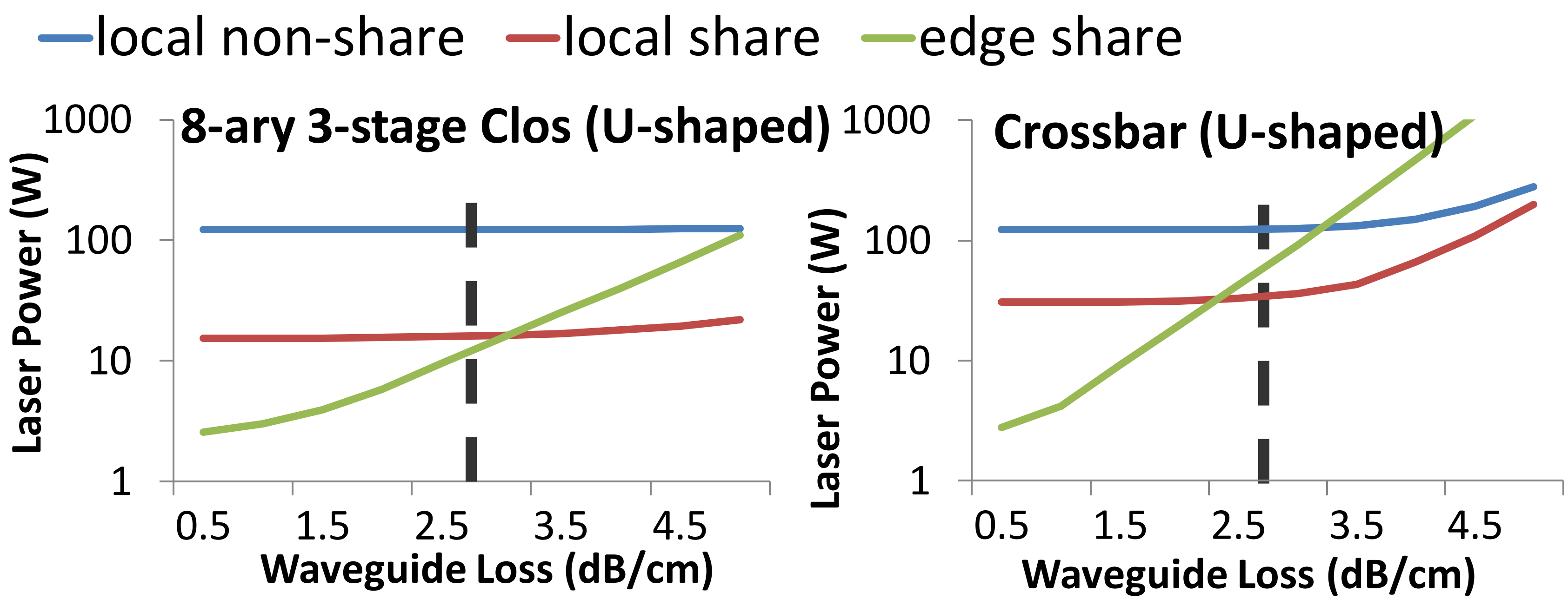


	Edge Placement	Local Placement
Sharing Degree	High	Low
Propagation loss	High	Low

Higher Sharing Degree
→ Higher η_{WPE}, Lower # of laser sources
→ Lower laser source power consumption

Higher Propagation loss
→ Higher required output optical power
→ Higher laser source power consumption

Laser Source Placement and Sharing



- For laser source power reduction, we need to ***jointly*** consider:
  - Network topology/layout;
  - Waveguide design;
  - Laser source sharing and placement.

Summary

- Cross-layer, thermally-aware optimizer for floorplanning of PNoCs
- Runtime workload allocation for thermal tuning power reduction
- Laser source placement for power/performance optimization
- ***Cross-layer simulation flow:*** an enabler to optimization of systems with PNoCs