Thermal Modeling and Analysis for Silicon Photonic Interconnect Networks

Jiang Xu

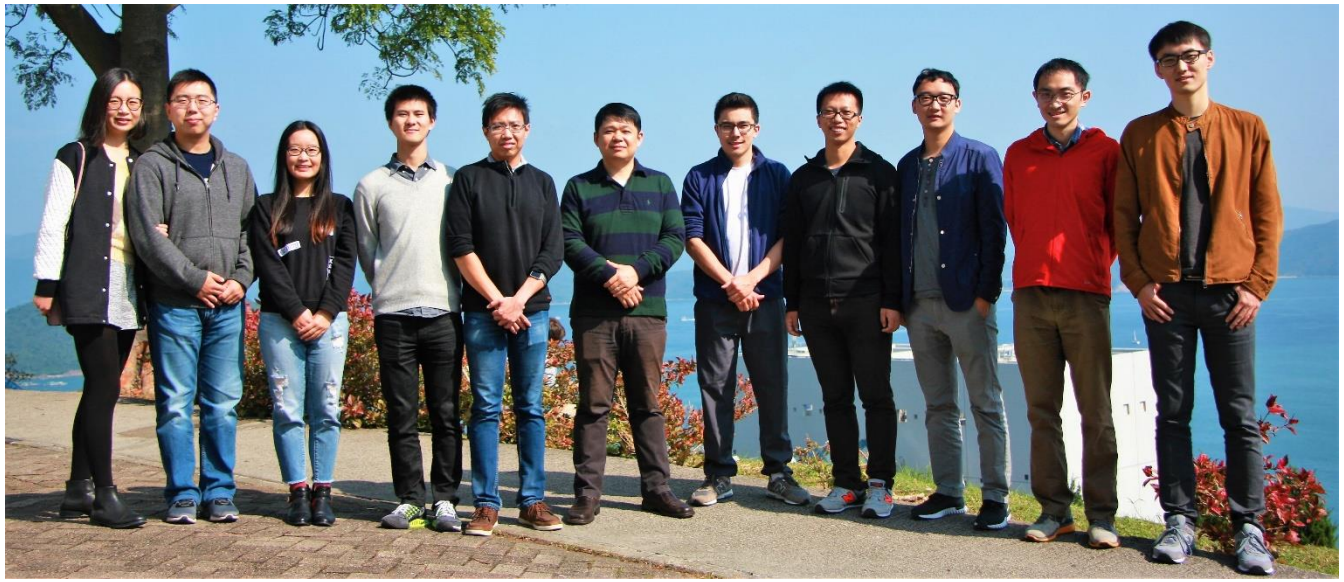
BIG DATA SYSTEM LAB

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

OPTICS Lab

Acknowledgement

- Current PhD students
  - Zhehui Wang, Duong Huu Kinh Luan, Peng Yang, Zhifei Wang, Zhe Wang, Haoran Li, Rafael Kioji Vivas Maeda, Xuanqi Chen, Zhongyuan Tian
- Past members
  - Mahdi Nikdast, Yaoyao Ye, Weichen Liu, Xiaowen Wu, Xuan Wang, Xing Wen, Kwai Hung Mo, Sébastien Le Beux, Yu Wang, Yiyuan Xie, Huaxi Gu



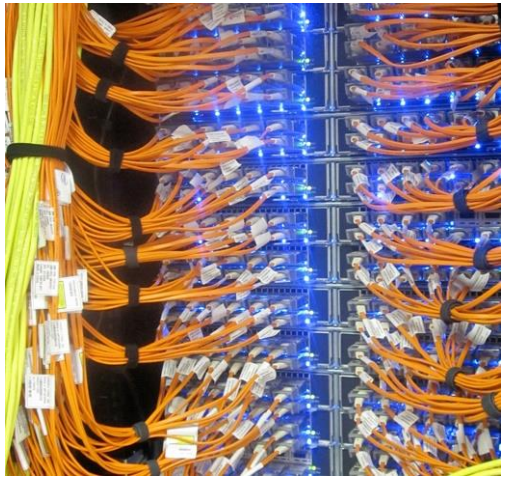
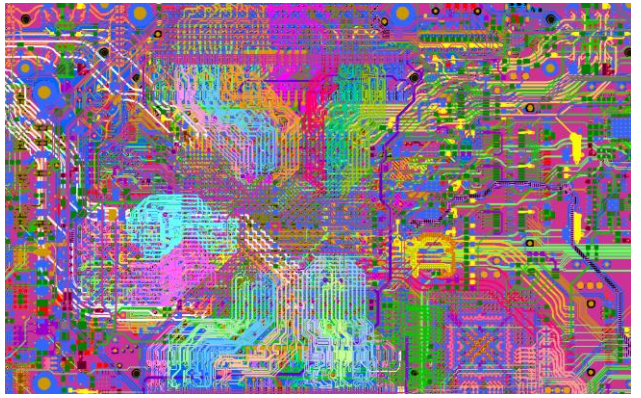
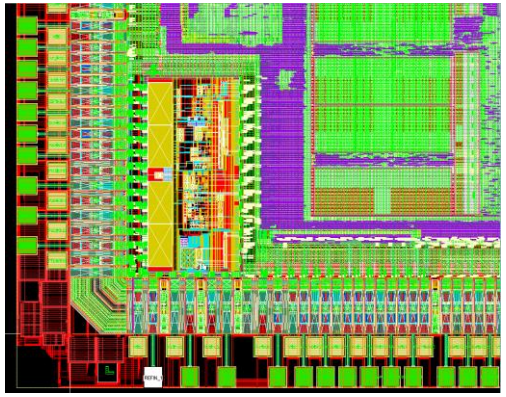
UGC 大學教育資助委員會 University Grants Committee

intel

XILINX

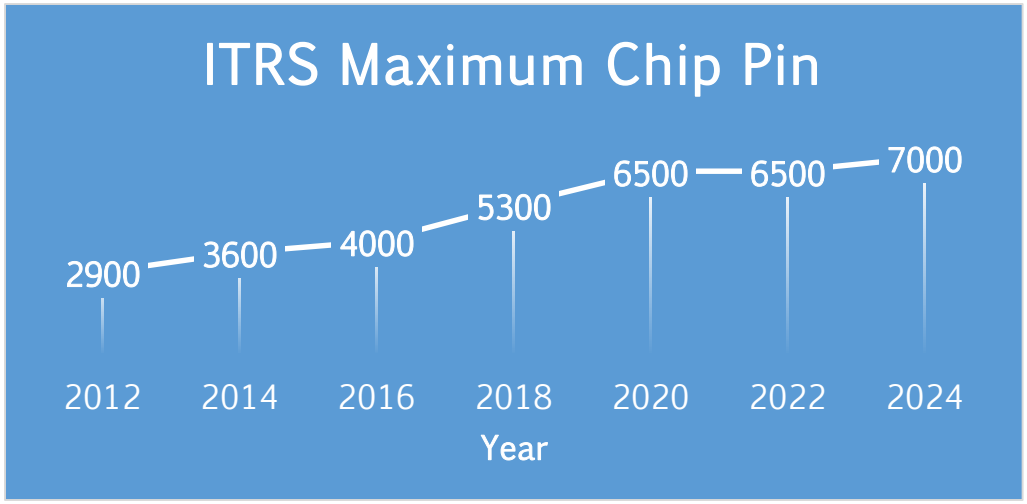
Interconnection Network: Road System of Computer

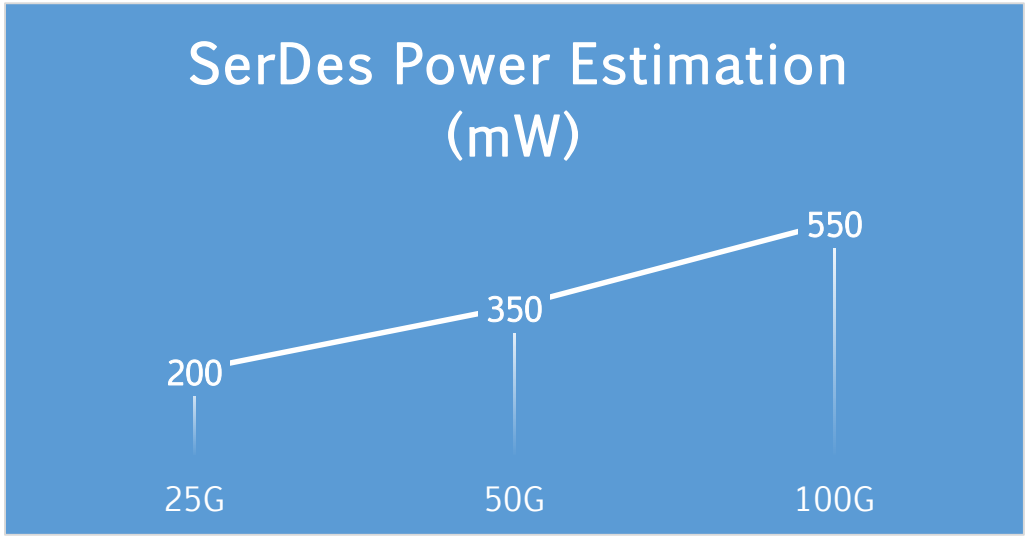
- WAN, LAN, SAN, rack, board, chip
- Key to system organization
  - Partition/tradeoff computation-communication-memory
  - Synchronization scheme
- Impact system performance
  - Cooperation among functional units
  - Communication latency and throughput
- Affect energy consumption
  - Significant energy is used to communicate
  - Idle power while waiting for information
- Decide system cost and chip yield
  - Cable, switch, and router
  - Chip metal layers, package pin, and PCB

Challenges of Electrical Interconnect

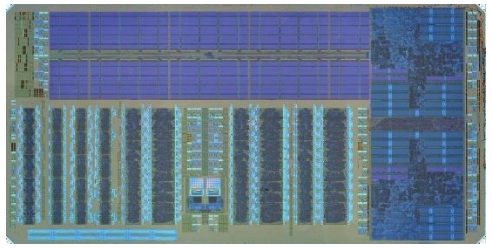
- More communications from more cores and memories
  - Blade server, micro server, disaggregated server …
  - Cisco QuantumFlow/40, Intel Phi/72, Tilera Tile/72, PicoChip/300 …
- Tighter I/O bandwidth
  - Maximum pin count of package grows slow
  - Higher packaging, PCB, and cabling cost
- Larger latency
  - Multiple clock cycles are required to cross a chip
  - Millions clock cycles to cross nodes in a rack
  - Billions clock cycles to cross racks
- Higher energy consumption and loss
  - Dynamic and leakage power of drivers and buffers
  - ~35dB/m @12.5G on high-quality PCB
- SerDes energy and performance bottleneck
  - ~5pJ/bit @ 100G





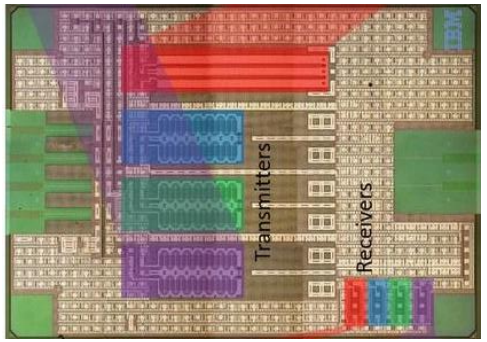
Silicon Photonics

- Successfully used in WAN and LAN
  - Multicomputer systems, Internet core routers, *etc.*
- Benefit from more matured silicon-based technologies and existing fabs
  - Micron-scale nanosecond-level devices are widely demonstrated
- Commercialization
  - IBM, Intel (Omni-Path), HP (Machine), Oracle (UNIC), Cisco, Mellanox, ST, NTT, NEC, Fujitsu (PECST), Huawei, ZTE …
  - Startups: Luxtera-ST, Lightwire/Cisco, Kotura/Mellanox, Caliopa/Huawei, Aurrion/Juniper, OneChip, Skorpios, Ayar…
  - PEDA: Cadence-PhoeniX-Lumerical, Mentor Graphics-Lumerical, RSoft/Synopsys …
- Questions remain
  - What difference can silicon photonics make?
  - How to maximize its benefits?
  - WHY?



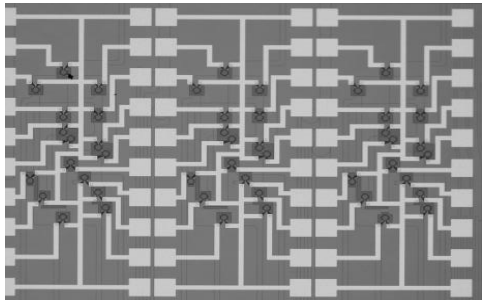
Integrated OE Interfaces & Processor

C. Sun *et al.*, "Single-chip microprocessor that communicates directly using light", Nature 2015



Integrated OE Interfaces

D.M. Gill, *et al.*, "Demonstration of Error Free Operation Up To 32 Gb/s From a CMOS Integrated Monolithic Nano-Photonic Transmitter", CLEO 2015



Integrated Optical Switches

R. Ji, *et al.* "Five-Port Optical Router Based on Microring Switches for Photonic Networks-on-Chip", IEEE Photonics Technology Letters 2013

Fundamentally Different "Building Material"

- Advantages
  - Ultra-high bandwidth
  - Low propagation delay
  - Low propagation loss
  - Low sensitivity to environmental EMI
- Disadvantages
  - Thermal sensitivity
  - Crosstalk noise
  - Process variation
  - Electrical/optical conversion
  - Difficult to "buffer"
- Differences bring challenges and new opportunities

Stone 85/220m
Solkan Bridge
Slovenia 1906

Steel 210/371m
Cold Spring Bridge
USA 1963

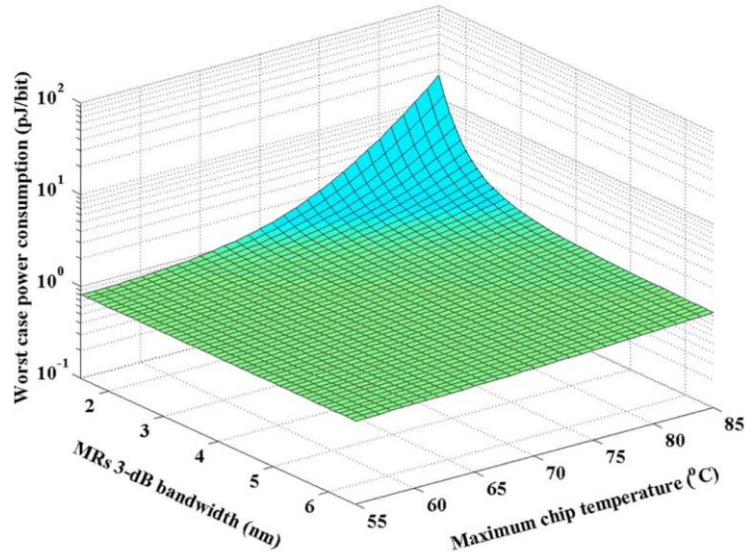
Steel 1377/2160m
Tsing Ma Bridge
Hong Kong 1997

Outline

- System-level optical thermal effect modeling and analysis
- Key findings and proposed techniques
- Case study
- Conclusions

Optical Thermal Effects

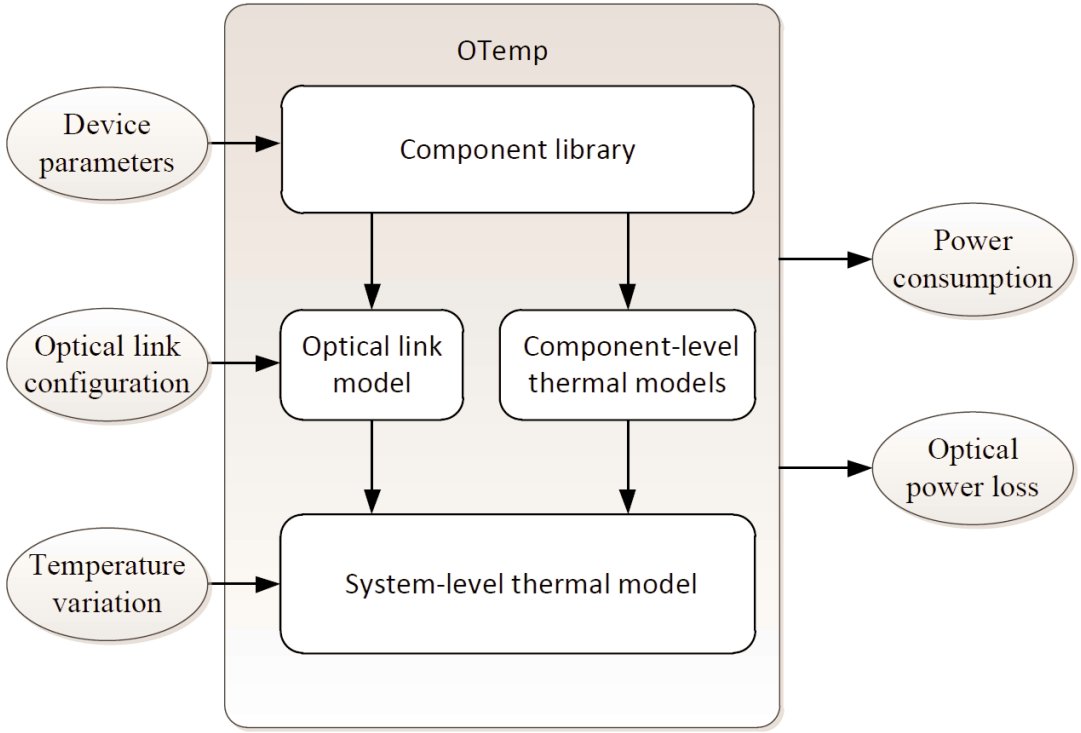
- Thermal sensitivity is a key issue of photonic devices
- Thermal effects can cause
  - Laser power efficiency degradation
  - Temperature-dependent wavelength shifting
  - Optical power loss due to wavelength mismatch
- System-level thermal model should consider
  - Temperature distribution
  - Waveguide propagation loss variation
  - Photodetector sensitivity and dark current
  - Laser temperature-dependent wavelength shifting and power efficiency
  - Device temperature-dependent wavelength shifting and optical power loss



* Y. Ye, *et al.*, "System-Level Modeling and Analysis of Thermal Effects in WDM-Based Optical Networks-on-Chip," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems 2014
* Y. Ye, *et al.*, "System-Level Modeling and Analysis of Thermal Effects in Optical Networks-on-Chip", IEEE Transactions on Very Large Scale Integration Systems 2013
* Y. Ye, *et al.*, "Modeling and Analysis of Thermal Effects in Optical Networks-on-Chip", ISVLSI 2011

OTEMP

- System-level optical thermal effect modeling and analysis platform
  - For inter-chip and intra-chip optical network, and optical switch
- Key design configurations
  - No. of WDM wavelengths and channel spacing
  - Electronic-based and thermal-based switching
  - Off-chip and on-chip lasers
  - Direct or BOME-based modulation
  - Active and parking BOSEs
  - Optical channel remapping
  - Temperature variations
- Open source www.ece.ust.hk/~eexu



Major Design Parameters

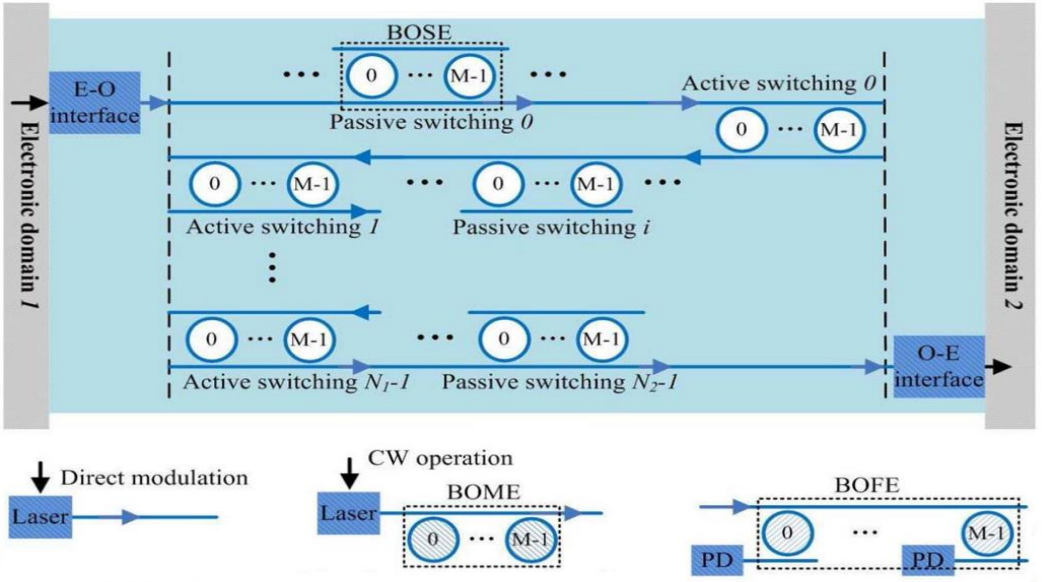
λ_{L_0}	laser wavelength at room temperature
ρ_L	laser wavelength shift coefficient
A	laser minimum threshold current
B	coefficient related to the laser threshold current
T_{th}	temperature at which threshold current is minimum
E	slope efficiency of laser at 0°C
Γ	coefficient related to the VCSEL slope efficiency
U_{slope}	slope of U-I laser characteristic curve
U_{th}	intercept of U-I laser characteristic curve
λ_{MR_0}	default initial setting of MRs

$\sigma_{fabrication}$	process variation
$\lambda_{MR\_optimal\_0}$	optimal initial setting of MRs
$L_{MR\_resonance}$	insertion loss at resonance
$E_{deserializer}$	deserializer power
$E_{serializer}$	serializer power
$\Delta\lambda_{misplace\_factor}$	width of misplace region
$\Delta\lambda_{elec\_switch\_off\_on}$	blue-shift of electrical switching
$\Delta\lambda_{thermal\_switch\_off\_on}$	red-shift of thermal switching
$\Delta\lambda_{modulation\_0\_1}$	blue-shift of modulation
$P_{modulator\_data\_0}$	modulator output power

ρ_{MR}	wavelength shift coefficient
Q	quality factor of MRs
E_{driver}	laser driver power
E_{PD}	photodetector power
P_{MR_on}	on-state MR power
$L_{crossing}$	waveguide crossing loss
L_{WG}	WG propagation loss
S_{RX}	receiver sensitivity
$P_{thermal}$	thermal adjustment power
E_{TIA_LA}	TIA-LA power

Link-based WDM Optical Interconnect Model

- Optical networks are composed of optical links
- An optical link includes
  - Laser source
  - Basic optical modulation element (BOME)
  - Basic optical switching element (BOSE)
  - Basic optical filter element (BOFE)
  - Photodetector (PD)



- Necessary condition for functional optical links

$$10\log\left(\left(I-\alpha-\beta\left(T_{\mathrm{L}}-T_{\mathrm{th}}\right)^2\right)\left(\varepsilon-\gamma\cdot T_{\mathrm{L}}\right)\right)-L_{\mathrm{BOME\_}x}-\sum_{i=0}^{N_1-1}L_{\mathrm{BOSE\_active\_}k}-\sum_{j=0}^{N_2-1}L_{\mathrm{BOSE\_parking\_}j}-L_{\mathrm{BOFE\_}x}-L_{\mathrm{WG}}\geq S_{\mathrm{RX}}$$

  - Optical power reaching PD must be larger than PD sensitivity

Laser Thermal Modeling

- Emission wavelength λ_L

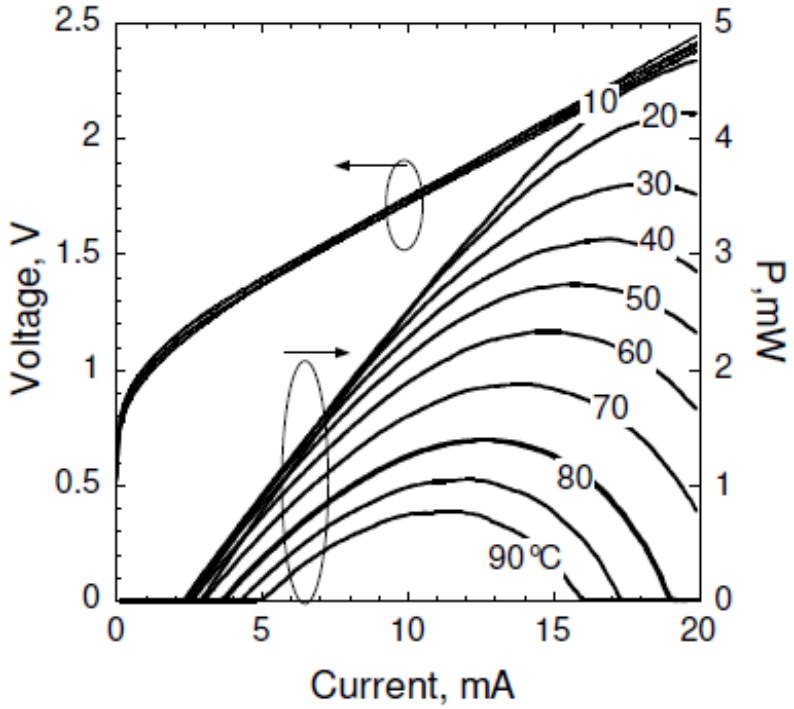
$$l_L \cdot n_{ave} = m_L \cdot \lambda_L / 2$$

- Temperature-dependent wavelength shift

$$\lambda_L = \lambda_{L\_0} + \rho_L (T_L - T_0)$$

- Output power under temperature T_L

$$P_{out} = (I - \alpha - \beta (T_L - T_{th})^2)(\varepsilon - \gamma \cdot T_L)$$



*Syrbu, OFC/NFOEC'08

- On-chip laser source, T_L varies over the on-chip temperature range
- Off-chip laser source with temperature control, T_L is fixed

BOME Thermal Model

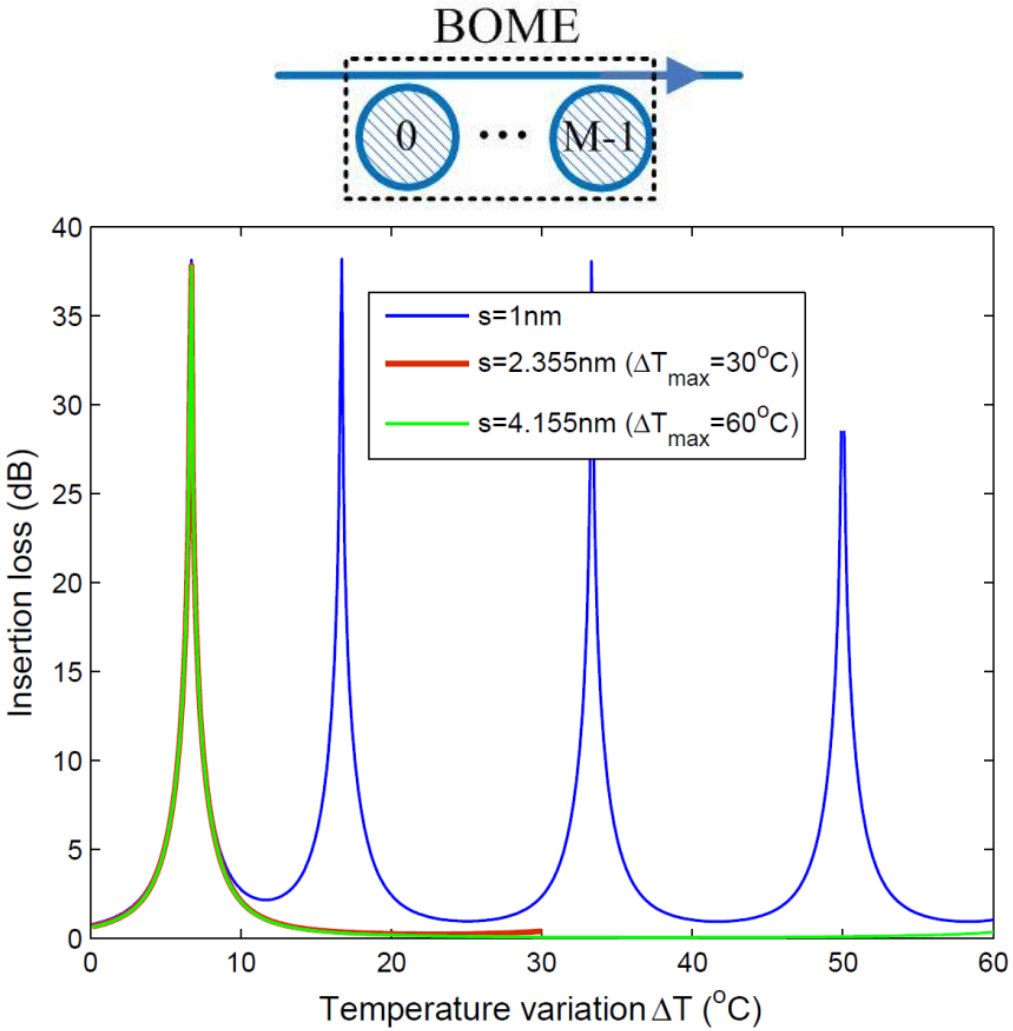
- Worst-case insertion loss to λ_0 under temperature variation ΔT

$$L_{BOME\_0} = \sum_{i=0}^{M-1} 10log \frac{(\frac{i \cdot s - b + \rho_{MR} \cdot \Delta T}{\delta})^2 + 1}{(\frac{i \cdot s - b + \rho_{MR} \cdot \Delta T}{\delta})^2 + (\frac{\kappa^2 - \kappa_p^2}{\kappa^2 + \kappa_p^2})^2}$$

- Worst-case insertion loss to λ_x under temperature variation ΔT

$$L_{BOME\_x} = \sum_{i=0}^{M-x-1} 10log \frac{(\frac{i \cdot s - b + \rho_{MR} \cdot \Delta T}{\delta})^2 + 1}{(\frac{i \cdot s - b + \rho_{MR} \cdot \Delta T}{\delta})^2 + (\frac{\kappa^2 - \kappa_p^2}{\kappa^2 + \kappa_p^2})^2} + \sum_{j=1}^{x} 10log \frac{(\frac{j \cdot s - \rho_{MR} \cdot \Delta T}{\delta})^2 + 1}{(\frac{j \cdot s - \rho_{MR} \cdot \Delta T}{\delta})^2 + (\frac{\kappa^2 - \kappa_p^2}{\kappa^2 + \kappa_p^2})^2}$$

- Findings
  - Large channel spacing can reduce thermal effects
  - Except for small temperature variations



Electrical-switched BOME on λ_7, Q=5000

BOSE Thermal Model

- For active switching, *M*-wavelength BOSE insertion loss to optical signal ω_n

$$L_{BOSE\_active}(\omega_n) = -10\log|f_{M-1}(\omega_n)|^2$$

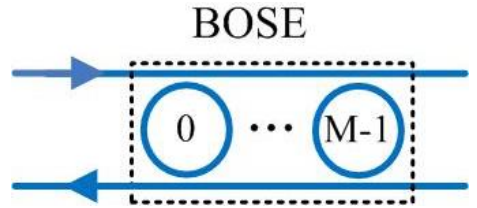
$$f_m(\omega_n) = r_m - \frac{t_m^i t_m^o}{r_m - f_{m-1}^{-1} exp(j2\theta_{m-1})} \qquad m = 1, 2, \ldots, M-1$$

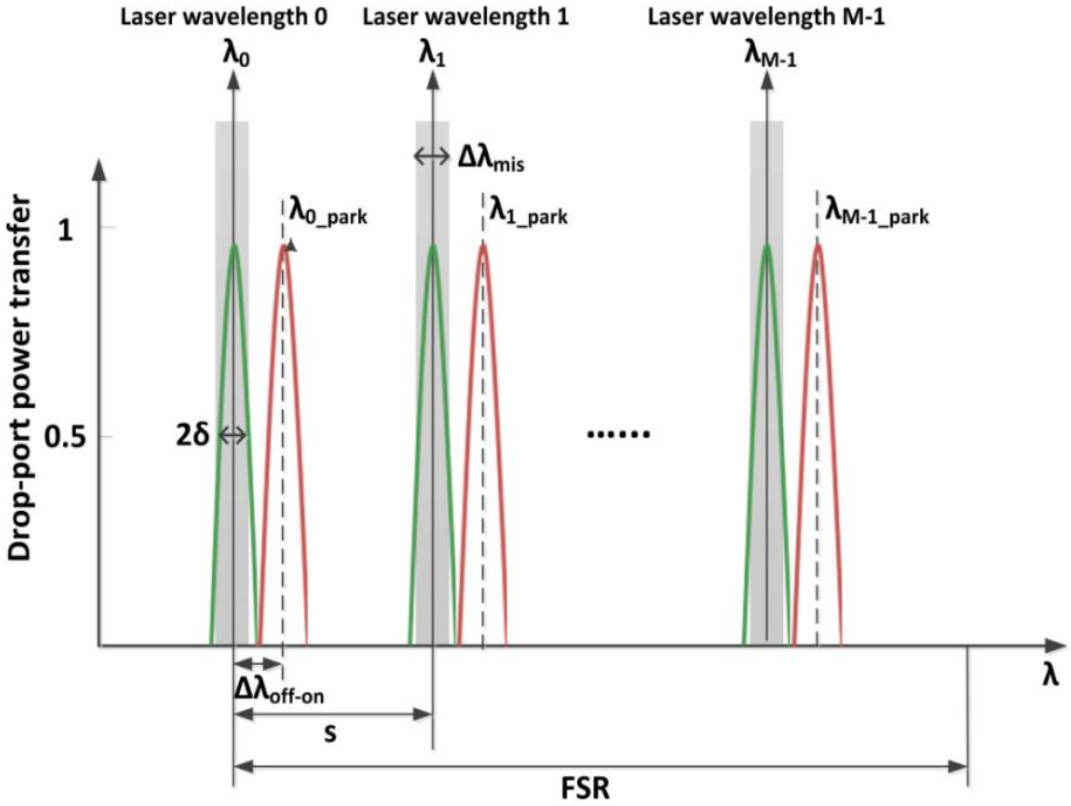
$$r_m = \frac{2\kappa^2}{j2\tau(\omega - \omega_m) + (2\kappa^2 + \kappa_p^2)} \qquad m = 1, 2, \ldots, M-1$$

$$t_m^i = t_m^o = \frac{j2\tau(\omega_n - \omega_m) + \kappa_p^2}{j2\tau(\omega_n - \omega_m) + (2\kappa^2 + \kappa_p^2)} \qquad m = 1, 2, \ldots, M-1$$

$$f_0(\omega_n) = r_0$$

$$\omega_m = \frac{2\pi c}{\lambda_m + \rho_{MR} \cdot \Delta T} \qquad m = 0, 1, 2, \ldots, M-1$$

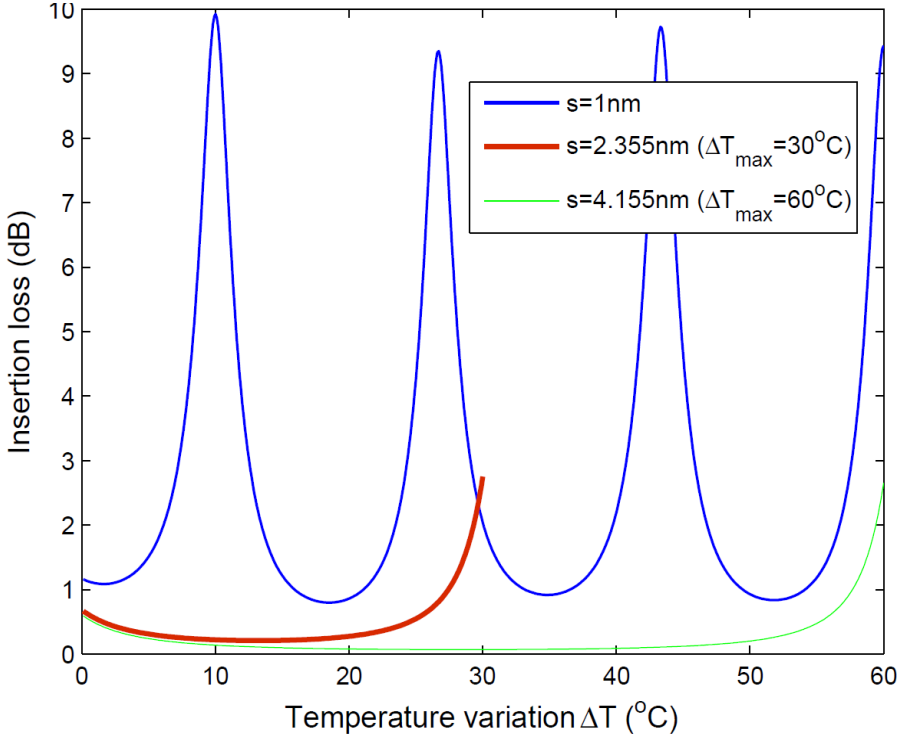




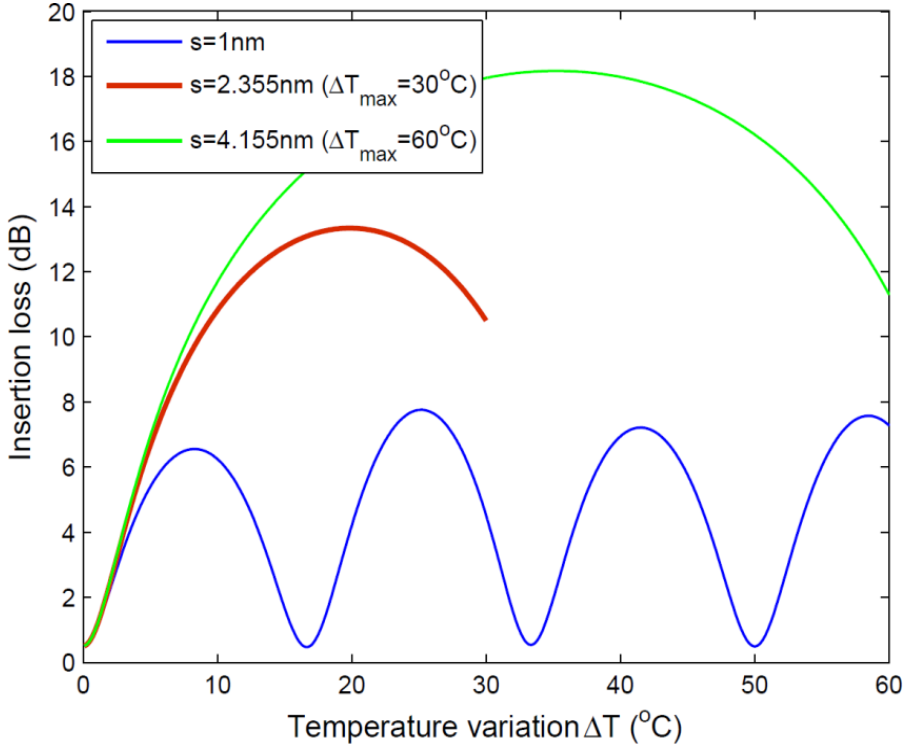
Active and parking BOSE

BOSE Thermal Model

- Large channel spacing
  - Help parking BOSE against thermal variations
  - Worsen active BOSE under thermal variations



Insertion loss of parking BOSE (8 wavelengths, Q=5000)



Insertion loss of active BOSE (8 wavelengths, Q=5000)

BOFE Thermal Model

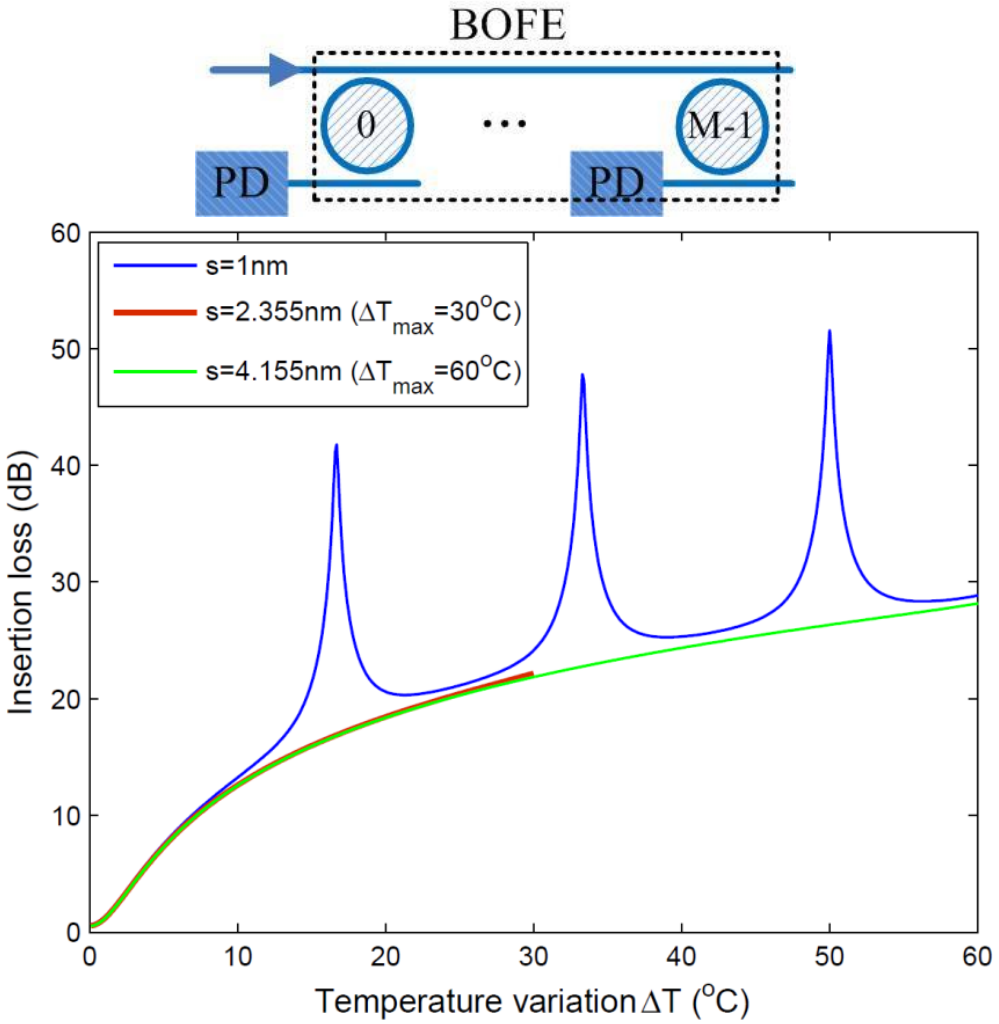
- BOFE insertion loss to wavelength λ_0 under temperature variation ΔT

$$L_{BOFE\_0} = 10log((\frac{2\kappa^2 + \kappa_p^2}{2\kappa^2})^2 \cdot (\frac{(\rho_{MR}\Delta T)^2 + \delta^2}{\delta^2}))$$

- BOFE insertion loss to wavelength λ_x under temperature variation ΔT

$$L_{BOFE\_x} = 10log((\frac{2\kappa^2 + \kappa_p^2}{2\kappa^2})^2 \cdot (\frac{(\rho_{MR}\Delta T)^2 + \delta^2}{\delta^2})) + \sum_{i=0}^{x-1} 10log\frac{((x-i)\cdot s + \rho_{MR}\Delta T)^2 + \delta^2}{((x-i)\cdot s + \rho_{MR}\Delta T)^2 + \delta^2 \cdot (\frac{\kappa_p^2}{2\kappa^2+\kappa_p^2})^2}$$

- Findings
  - Large channel spacing can reduce thermal effects
  - But still as high as 20dB for ΔT=30°C



Insertion loss of an 8-wavelength BOFE to λ_7, Q=5000

Waveguide and Receiver

- Waveguide propagation loss is proportional to the refractive index difference Φ between the core and cladding
  - Propagation loss under temperature variation ΔT

$$L_{\mathrm{WG}} = L_{\mathrm{WG\_0}} \cdot \left(1 + \frac{\sigma_c - \sigma_d}{\phi} \cdot \Delta T + \frac{(\sigma_c - \sigma_d)^2}{3\phi^2} \cdot (\Delta T)^2\right)$$

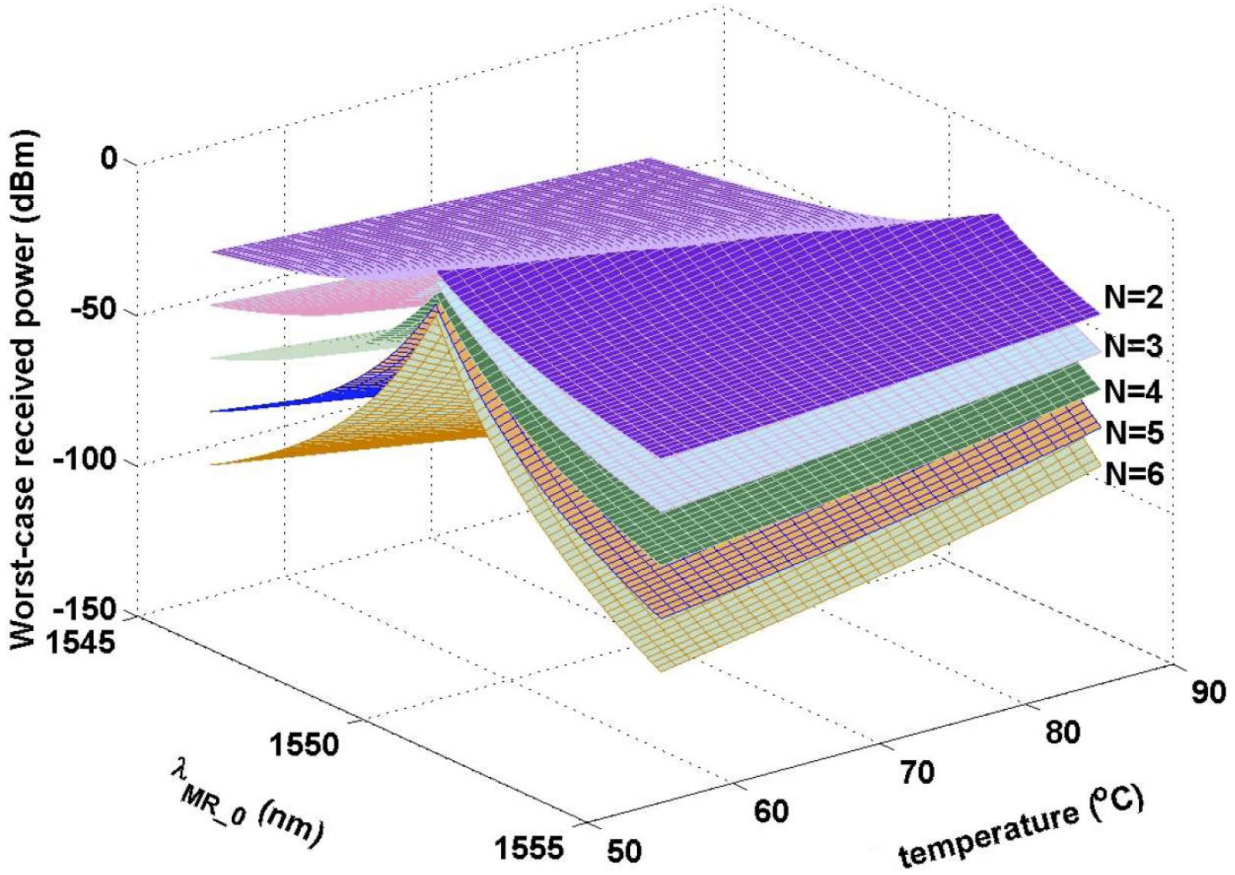
  - Loss variation on a Si/SiO_2 waveguide is about 0.22% for $\Delta T = 30^{o}C$
- Detection sensitivity of photodetector does not change obviously at high temperatures [Koester GFP'06]

Key Findings

- Optimal initial device settings
  - Regardless of network architectures
  - Minimize power consumption

$$\lambda_{MR\_i} = \lambda_{L\_i} + \frac{\rho_L - \rho_{MR}}{2} \cdot (T_{max} + T_{min} - 2T_0)$$

- Reducing BOME/BOSE/BOFE stages
  - Can significantly lower worst-case power consumption under thermal variations
  - Direct modulated laser does not need BOME and hence preferred



N is the number of stages

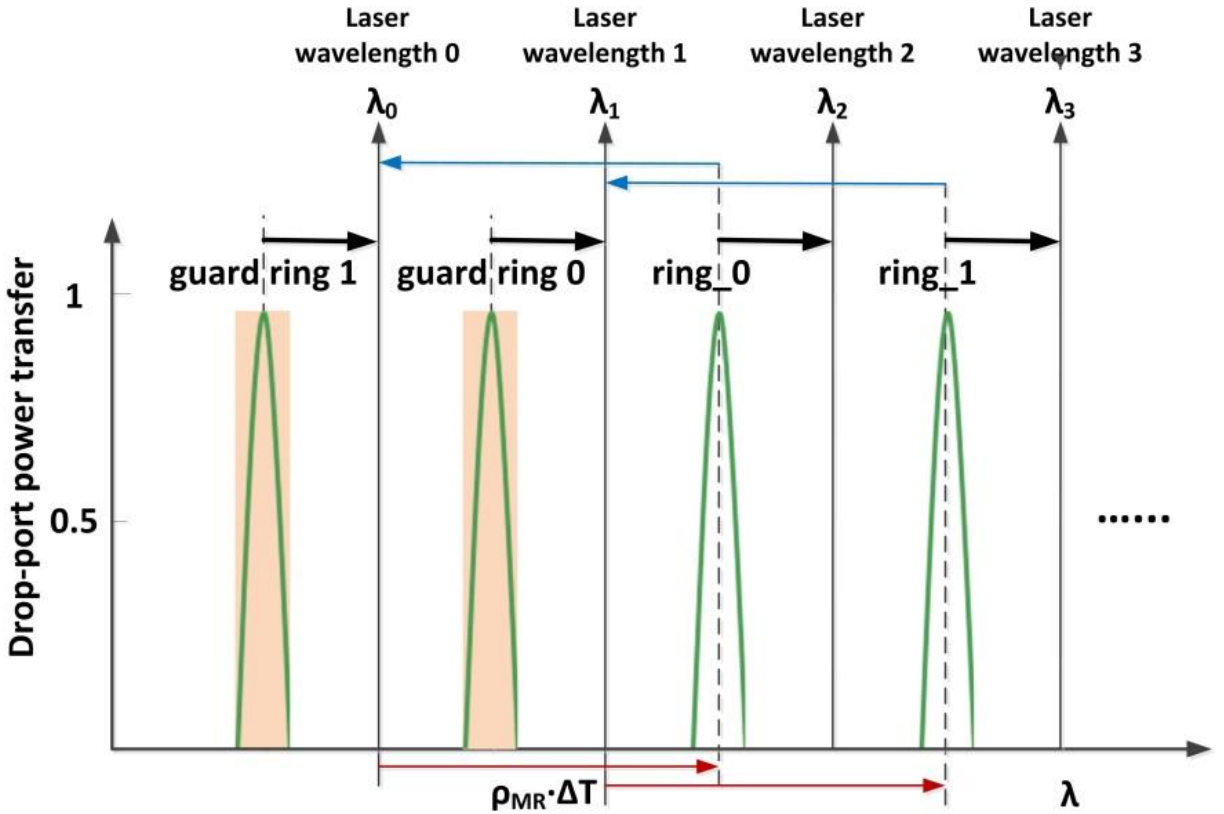
Channel Remapping with Guard MRs

- Remap channels to different MRs based on thermal variations
- Reduce tuning distance and save tuning power
- Tuning distance
  - Without channel remapping

$$d = \rho_{MR} \cdot (\Delta T_{max} - \Delta T)$$

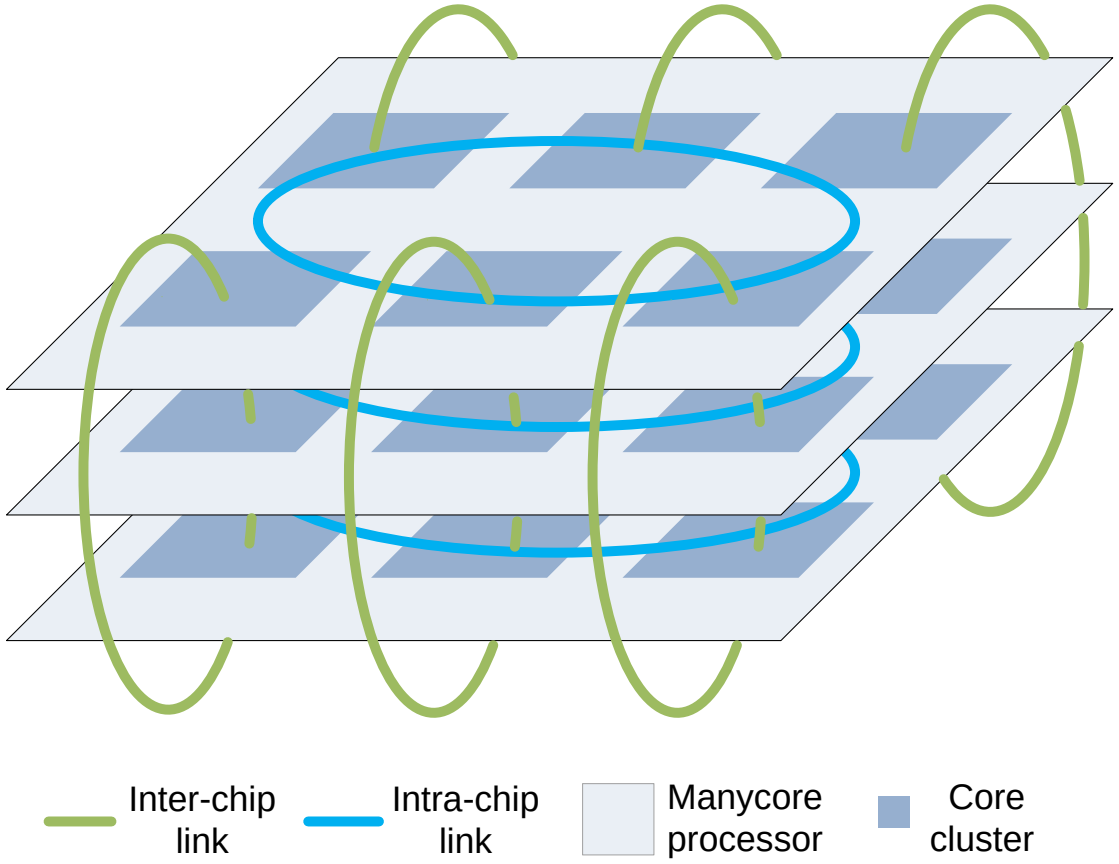
  - With channel remapping

$$d = \lceil \frac{\rho_{MR} \cdot \Delta T}{s} \rceil \cdot s - \rho_{MR} \cdot \Delta T$$

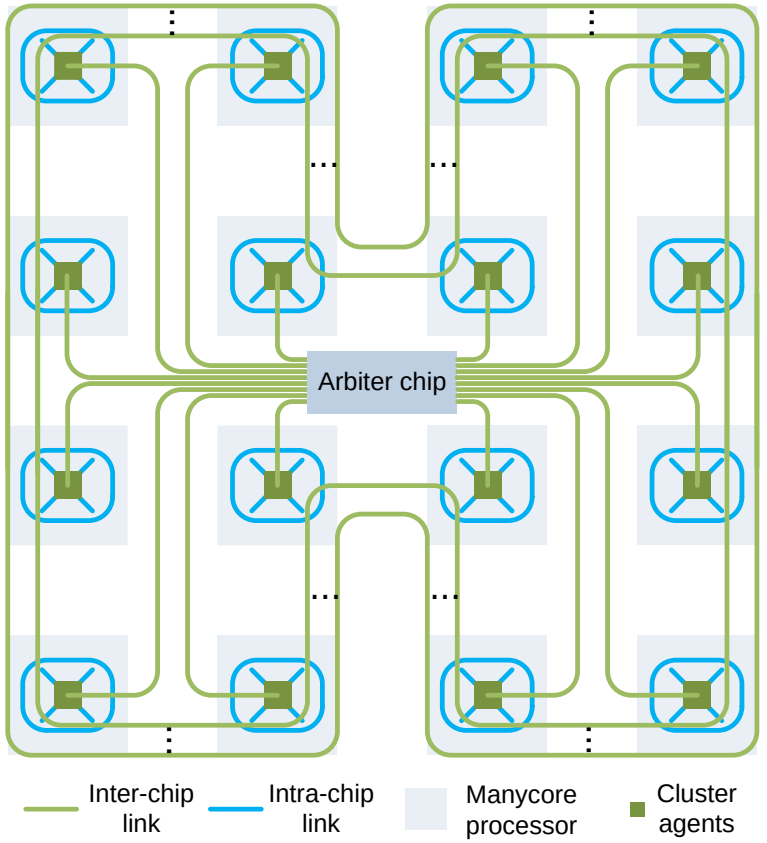


I²CON: Ring-Based Inter/Intra-Chip Optical Network

Logical view

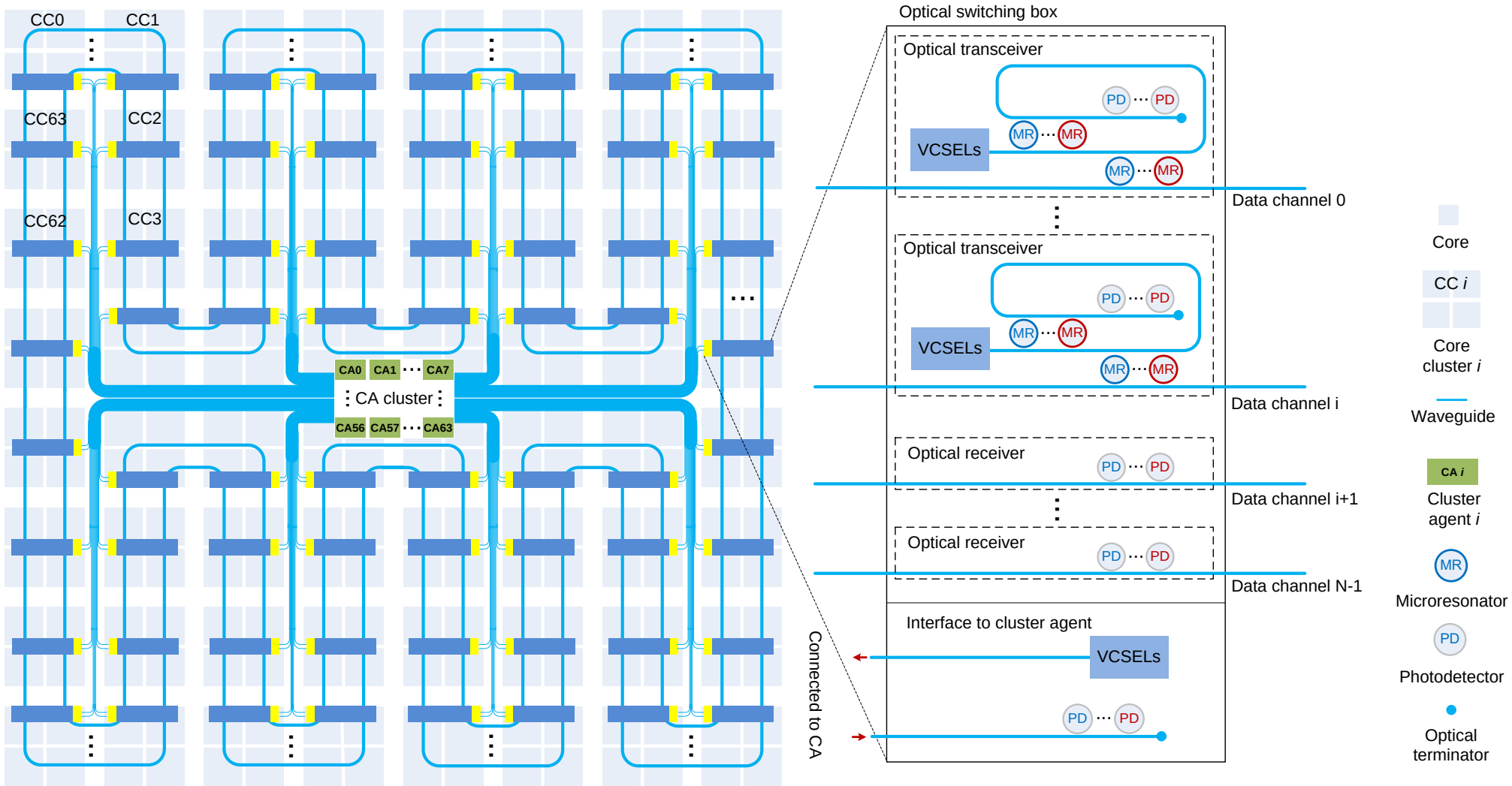


Multi-chip floorplan



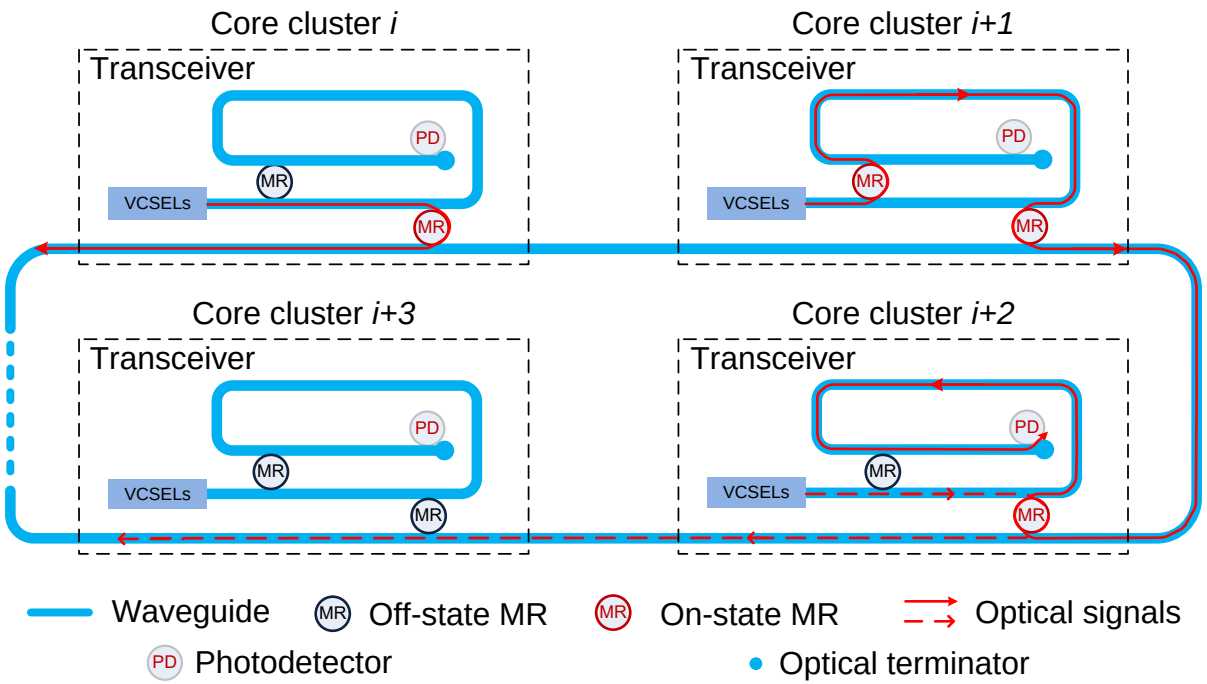
* X. Wu, *et al.* “SUOR: Sectioned Undirectional Optical Ring for Chip Multiprocessor,” ACM Journal on Emerging Technologies in Computing Systems 2014
* X. Wu, *et al.*, “An Inter/Intra-chip Optical Network for Manycore Processors" IEEE Transactions on Very Large Scale Integration Systems 2015

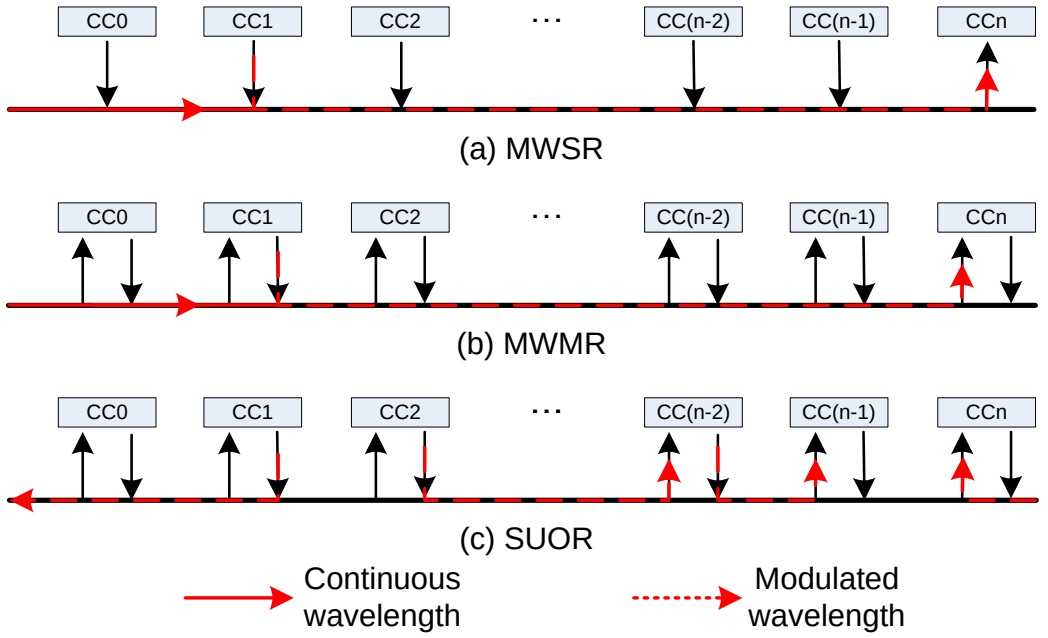
Intra-Chip Network of I²CON



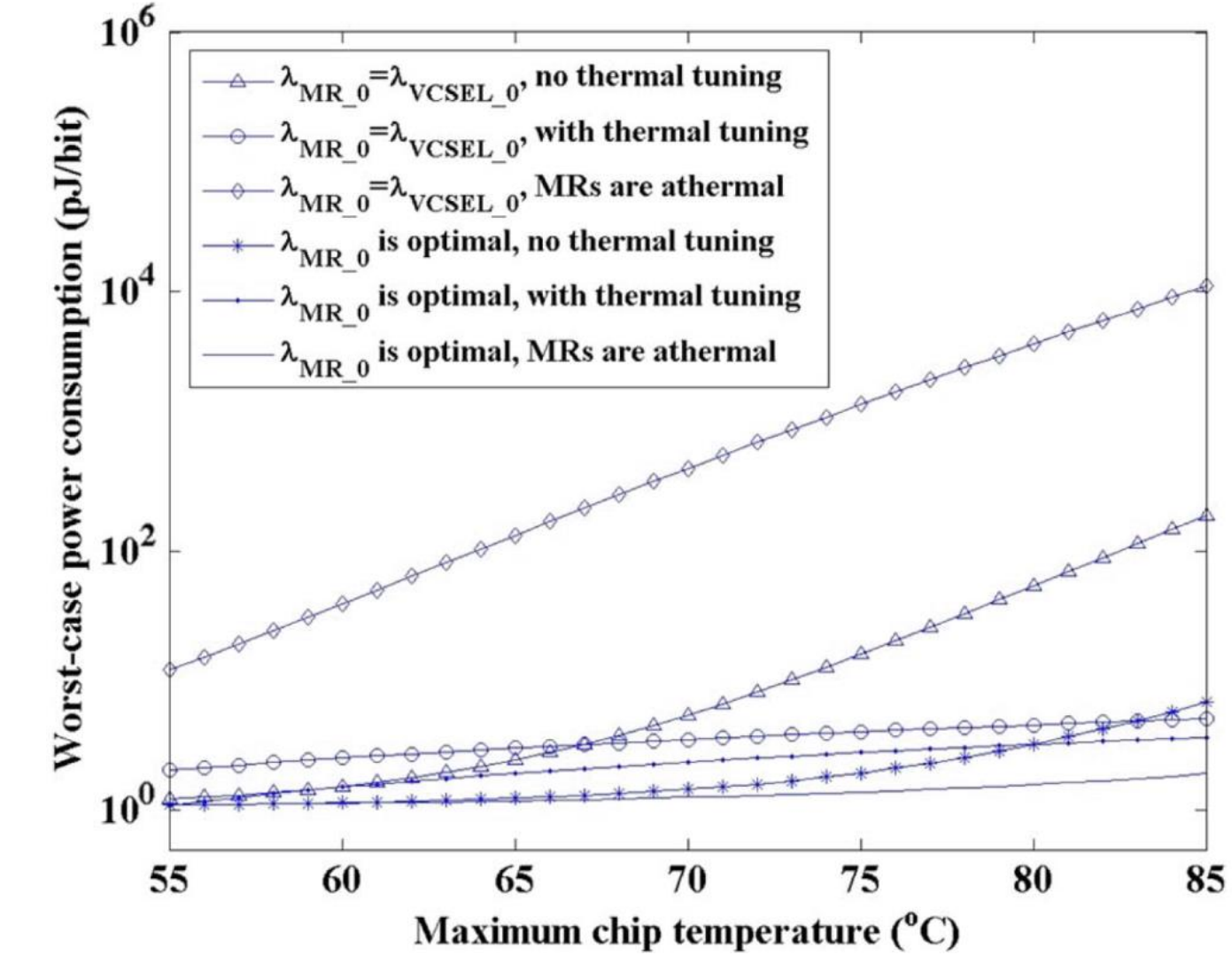
Segmented and Bidirectional Data Channels

- Improve network capacity and utilization
- Improve energy efficiency





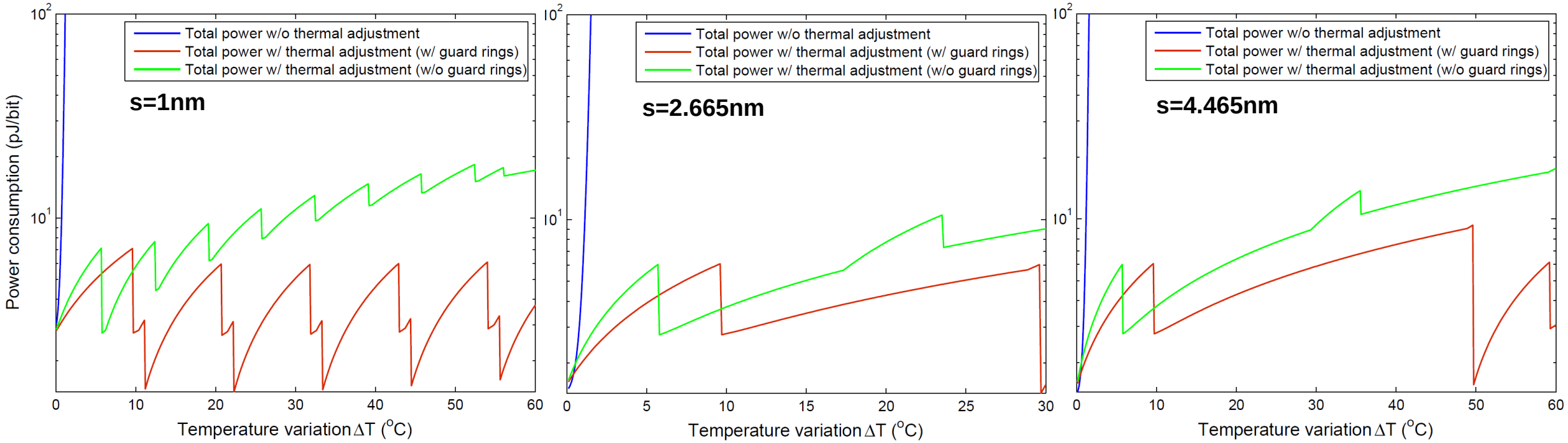
Optimal Initial Device Setting



- Improve the energy efficiency by 29%
  - Default setting is λ_{MR_0}= λ_{laser_0}=1550nm at room temperature
  - 3-dB bandwidth is 3.1nm,
  - Three switching stages

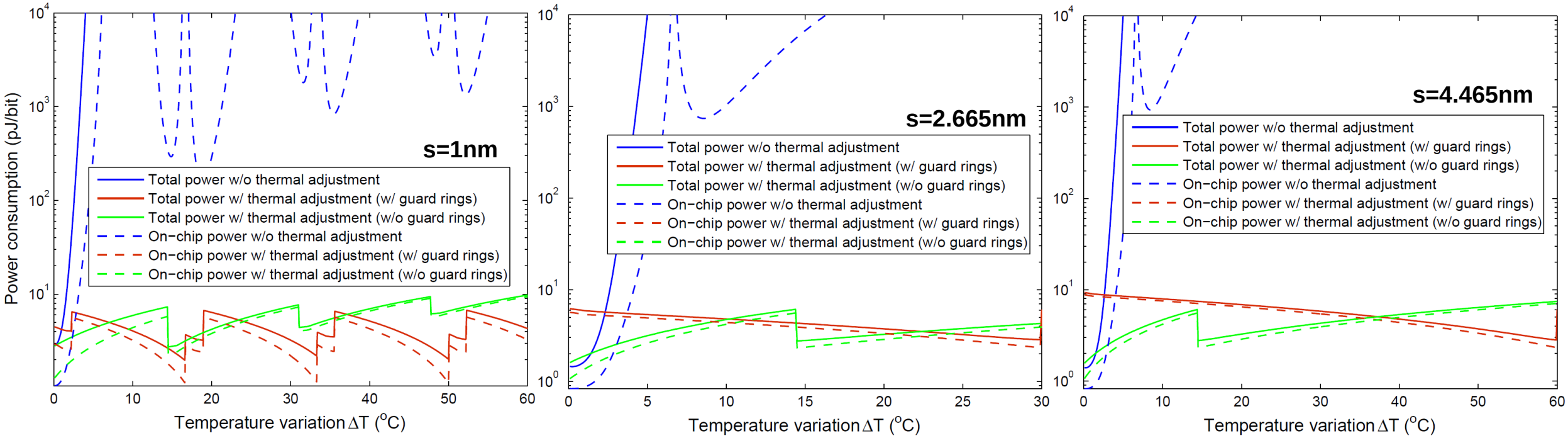
Using On-Chip Lasers

- Worst-case power consumption under maximum variation ΔT
  - 8 wavelengths, Q=5000, 3 active and 10 parking BOSEs
- Large channel spacing does not help

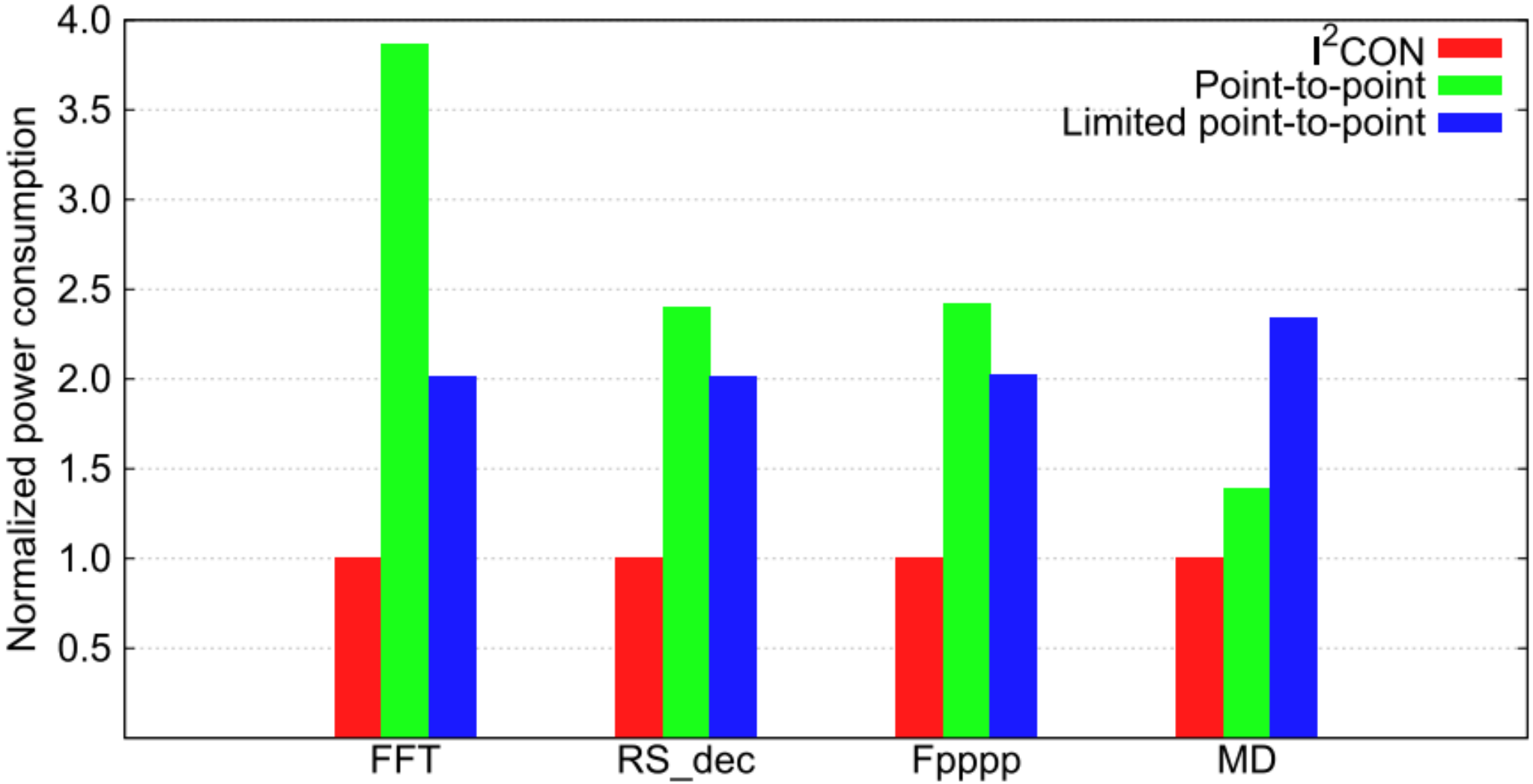


Using Off-Chip Lasers

- Worst-case power consumption under maximum variation ΔT
  - 8 wavelengths, Q=5000, 3 active and 10 parking BOSEs
- Large channel spacing does not help



Overall Power Consumption under Real Applications



Summary

- Systematically modeled and analyzed thermal effects at system level
- OTemp is developed and released
- Key findings
  - Optimal initial device settings
  - Lower switching stages
  - Channel remapping and guard MRs
  - Large channel spacing is not effective

Past

Present

Future

Publically Released Tools

- Bibliography for inter/intra-chip optical networks
- JADE heterogeneous multiprocessor simulation environment
- COSMIC heterogeneous multiprocessor benchmark suite
- CLAP optical crosstalk and loss analysis platform
- OTemp optical thermal effect modeling platform
- OEIL optical and electrical interface and link analysis environment
- MCSL realistic network-on-chip traffic patterns
- PowerSoC power delivery system analysis platform

www.ece.ust.hk/~eexu

Reference

1. Zhehui Wang, Jiang Xu, et al. “A Holistic Modelling and Analysis of Optical-Electrical Interfaces for Inter/Intra-chip Interconnects,” IEEE Transactions on Very Large Scale Integration Systems, July 2016.
2. Peng Yang, Shigeru Nakamura, Kenichiro Yashiki, Zhehui Wang, Luan H. K. Duong, Zhifei Wang, Xuanqi Chen, Yuichi Nakamura, Jiang Xu, “Inter/Intra-Chip Optical Interconnection Network: Opportunities, Challenges, and Implementations” IEEE/ACM International Symposium on Networks-on-Chip (NOCS), Japan, 2016.
3. Xiaowen Wu, Jiang Xu, et al., “An Inter/Intra-chip Optical Network for Manycore Processors," IEEE Transactions on Very Large Scale Integration Systems, vol.23, no.4, pp.678-691, April 2015.
4. Yaoyao Ye, Zhehui Wang, Peng Yang, Jiang Xu, et al. “System-Level Modeling and Analysis of Thermal Effects in WDM-Based Optical Networks-on-Chip,” IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 33, no. 11, pp. 1718-1731, November 2014.
5. Xiaowen Wu, Jiang Xu, et al. “SUOR: Sectioned Undirectional Optical Ring for Chip Multiprocessor,” ACM Journal on Emerging Technologies in Computing Systems, vol. 10 no. 4, May 2014.
6. Yaoyao Ye, Jiang Xu, et al. “System-Level Modeling and Analysis of Thermal Effects in Optical Networks-on-Chip”, IEEE Transactions on Very Large Scale Integration Systems, February 2013.
7. Yaoyao Ye, Jiang Xu, et al. “Thermal Analysis for 3D Optical Network-on-Chip Based on a Novel Low-Cost 6x6 Optical Router”, IEEE Optical Interconnects Conference, May 2012.
8. Yaoyao Ye, Jiang Xu, et al. “Modeling and Analysis of Thermal Effects in Optical Networks-on-Chip”, in Proceedings of IEEE Computer Society Annual Symposium on VLSI, July 2011.