Thermal Modeling and Analysis for Silicon Photonic Interconnect Networks

Jiang Xu
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
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-PhoniX-Lumerical, Mentor Graphics-Lumerical, RSoft/Synopsys ...
- Questions remain
  - What difference can silicon photonics make?
  - How to maximize its benefits?
  - WHY?
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

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

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_{L,0}$</td>
<td>laser wavelength at room temperature</td>
</tr>
<tr>
<td>$\rho_L$</td>
<td>laser wavelength shift coefficient</td>
</tr>
<tr>
<td>$A$</td>
<td>laser minimum threshold current</td>
</tr>
<tr>
<td>$B$</td>
<td>coefficient related to the laser threshold current</td>
</tr>
<tr>
<td>$T_{th}$</td>
<td>temperature at which threshold current is minimum</td>
</tr>
<tr>
<td>$E$</td>
<td>slope efficiency of laser at 0°C</td>
</tr>
<tr>
<td>$\Gamma$</td>
<td>coefficient related to the VCSEL slope efficiency</td>
</tr>
<tr>
<td>$U_{slope}$</td>
<td>slope of U-I laser characteristic curve</td>
</tr>
<tr>
<td>$U_{th}$</td>
<td>intercept of U-I laser characteristic curve</td>
</tr>
<tr>
<td>$\lambda_{MR,0}$</td>
<td>default initial setting of MRs</td>
</tr>
<tr>
<td>$\sigma_{fabrication}$</td>
<td>process variation</td>
</tr>
<tr>
<td>$\lambda_{MR_{optimal,0}}$</td>
<td>optimal initial setting of MRs</td>
</tr>
<tr>
<td>$L_{MR_{resonance}}$</td>
<td>insertion loss at resonance</td>
</tr>
<tr>
<td>$E_{deserializer}$</td>
<td>deserializer power</td>
</tr>
<tr>
<td>$E_{serializer}$</td>
<td>serializer power</td>
</tr>
<tr>
<td>$\Delta\lambda_{misplace_factor}$</td>
<td>width of misplace region</td>
</tr>
<tr>
<td>$\Delta\lambda_{elec_switch_off_on}$</td>
<td>blue-shift of electrical switching</td>
</tr>
<tr>
<td>$\Delta\lambda_{thermal_switch_off_on}$</td>
<td>red-shift of thermal switching</td>
</tr>
<tr>
<td>$\Delta\lambda_{modulation_0_1}$</td>
<td>blue-shift of modulation</td>
</tr>
<tr>
<td>$P_{modulator_data_0}$</td>
<td>modulator output power</td>
</tr>
<tr>
<td>$\rho_{MR}$</td>
<td>wavelength shift coefficient</td>
</tr>
<tr>
<td>$Q$</td>
<td>quality factor of MRs</td>
</tr>
<tr>
<td>$E_{driver}$</td>
<td>laser driver power</td>
</tr>
<tr>
<td>$E_{PD}$</td>
<td>photodetector power</td>
</tr>
<tr>
<td>$P_{MR_{on}}$</td>
<td>on-state MR power</td>
</tr>
<tr>
<td>$L_{crossing}$</td>
<td>waveguide crossing loss</td>
</tr>
<tr>
<td>$L_{WG}$</td>
<td>WG propagation loss</td>
</tr>
<tr>
<td>$S_{RX}$</td>
<td>receiver sensitivity</td>
</tr>
<tr>
<td>$P_{thermal}$</td>
<td>thermal adjustment power</td>
</tr>
<tr>
<td>$E_{TIA_LA}$</td>
<td>TIA-LA power</td>
</tr>
</tbody>
</table>
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
  
  \[
  10 \log \left( \left( I - \alpha - \beta (T_L - T_{th})^2 \right) (\varepsilon - \gamma \cdot T_L) \right) - L_{BOME_x} - \sum_{i=0}^{N_1-1} L_{BOSE\_active_k} - \sum_{j=0}^{N_2-1} L_{BOSE\_parking_j} - L_{BOFE_x} - L_{WG} \geq S_{RX}
  \]

  - Optical power reaching PD must be larger than PD sensitivity
Laser Thermal Modeling

- Emission wavelength $\lambda_L$
  \[ l_L \cdot n_{ave} = m_L \cdot \lambda_L / 2 \]
- Temperature-dependent wavelength shift
  \[ \lambda_L = \lambda_{L0} + \rho_L (T_L - T_0) \]
- Output power under temperature $T_L$
  \[ P_{out} = (I - \alpha - \beta (T_L - T_{th})^2)(\varepsilon - \gamma \cdot T_L) \]
- On-chip laser source, $T_L$ varies over the on-chip temperature range
- Off-chip laser source with temperature control, $T_L$ is fixed

*Syrbu, OFC/NFOEC’08*
BOME Thermal Model

- Worst-case insertion loss to $\lambda_0$ under temperature variation $\Delta T$

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

- Worst-case insertion loss to $\lambda_x$ under temperature variation $\Delta T$

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

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

Electrical-switched BOME on $\lambda_7$, $Q=5000$
BOSE Thermal Model

- For active switching, $M$-wavelength BOSE insertion loss to optical signal $\omega_n$

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

$$f_m(\omega_n) = r_m - \frac{t^i_m t^0_m}{r_m - f^{-1}_{m-1} \exp(j2\theta_{m-1})}$$

$$r_m = \frac{2\kappa_p^2}{j2\tau(\omega - \omega_m) + (2\kappa_o^2 + \kappa_p^2)}$$

$$t^i_m = t^0_m = \frac{j2\tau(\omega_n - \omega_m) + \kappa_p^2}{j2\tau(\omega_n - \omega_m) + (2\kappa_o^2 + \kappa_p^2)}$$

$$f_0(\omega_n) = r_0$$

$$\omega_m = \frac{2\pi c}{\lambda_m + \rho_{MR} \cdot \Delta T}$$

$m = 1, 2, \ldots, M - 1$

$m = 1, 2, \ldots, M - 1$

$m = 1, 2, \ldots, M - 1$

$m = 0, 1, 2, \ldots, M - 1$
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 $\lambda_0$ under temperature variation $\Delta T$

$$L_{BOFE_0} = 10\log\left(\frac{2\kappa^2 + \kappa_p^2}{2\kappa^2}\cdot\left(\frac{\rho_{MR}\Delta T}{\delta^2} + \delta^2\right)\right)$$

- BOFE insertion loss to wavelength $\lambda_x$ under temperature variation $\Delta T$

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

- Findings
  - Large channel spacing can reduce thermal effects
  - But still as high as 20dB for $\Delta T=30^\circ C$
Waveguide and Receiver

- Waveguide propagation loss is proportional to the refractive index difference $\Phi$ between the core and cladding
  - Propagation loss under temperature variation $\Delta T$
    \[
    L_{WG} = L_{WG,0} \left(1 + \frac{\sigma_c - \sigma_d}{\Phi} \cdot \Delta T + \frac{(\sigma_c - \sigma_d)^2}{3\Phi^2} \cdot (\Delta T)^2\right)
    \]
  - Loss variation on a Si/SiO$_2$ waveguide is about 0.22% for $\Delta T=30^\circ C$
- 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
    \[ d = \left( \frac{\rho_{MR} \cdot \Delta T}{s} \right) \cdot s - \rho_{MR} \cdot \Delta T \]
I²CON: Ring-Based Inter/Intra-Chip Optical Network

Logical view

Multi-chip floorplan

* 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

Waveguide
Optical terminator
MR
Microresonator
PD
Photodetector
Core
Cluster agent
i
Optical switching box
Optical transceiver
VCSELs
PD
Photodetector
Interface to cluster agent
VCSELs
Optical transceiver
Data channel 0
Data channel i
Data channel i+1
Data channel N-1

Connected to CA
CA cluster
CC
Cluster agent
i
Waveguide
CA
i
Microresonator
PD
Photodetector
Optical terminator

2017-03-02
Jiang Xu (Big Data System Lab)
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 $\lambda_{MR,0} = \lambda_{laser,0} = 1550\text{nm}$ at room temperature
  - 3-dB bandwidth is 3.1nm,
  - Three switching stages
Using On-Chip Lasers

- Worst-case power consumption under maximum variation $\Delta T$
  - 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 $\Delta T$
  - 8 wavelengths, $Q=5000$, 3 active and 10 parking BOSEs

- Large channel spacing does not help
Overall Power Consumption under Real Applications

![Bar chart showing normalized power consumption for different applications: FFT, RS_dec, Fpppp, MD. The x-axis represents different applications, and the y-axis represents normalized power consumption. The chart compares different power consumption scenarios: I²CON, Point-to-point, Limited point-to-point.](chart.png)
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
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


