Silicon Photonic Interconnection Networks

Madeleine Glick

APIC Corporation

Collaborators

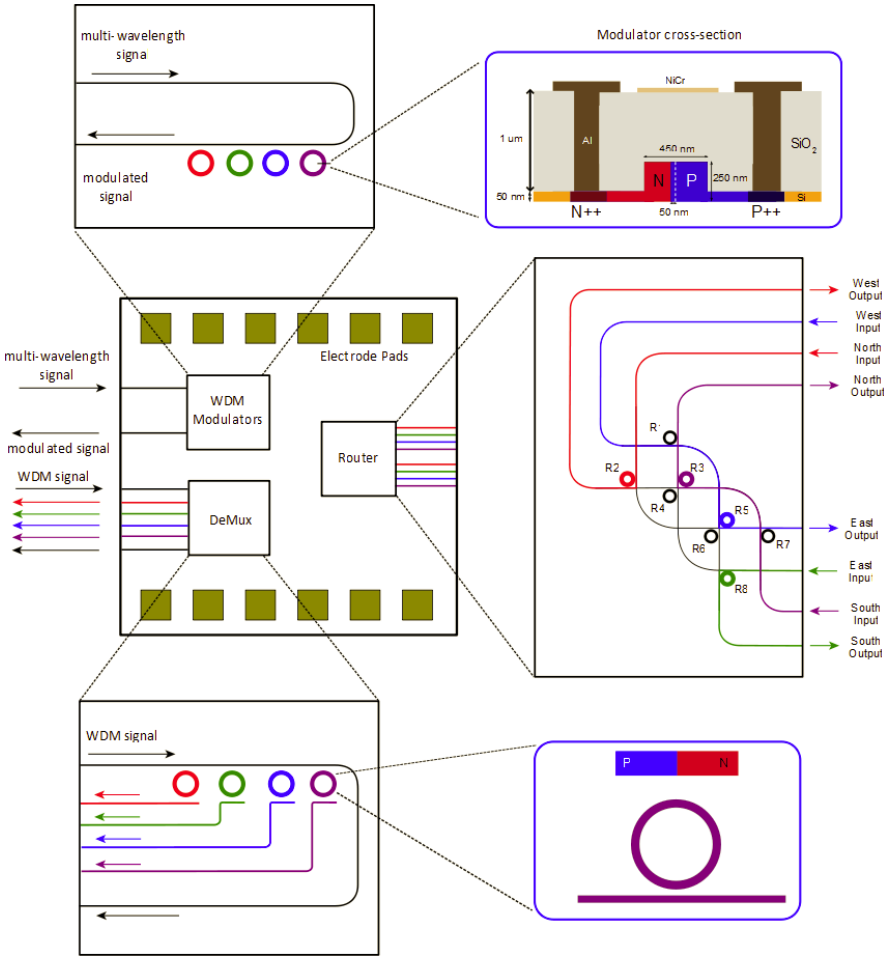
- Abhinav Rohit, Gerald Miller, Madeleine Glick, Raj Dutt

  APIC Corporation
- Sébastien Rumley, Qi Li, Gouri Dongaonkar, Noam Ophir, Kishore Padmaraju, Lee Zhu, David Calhourn, Keren Bergman

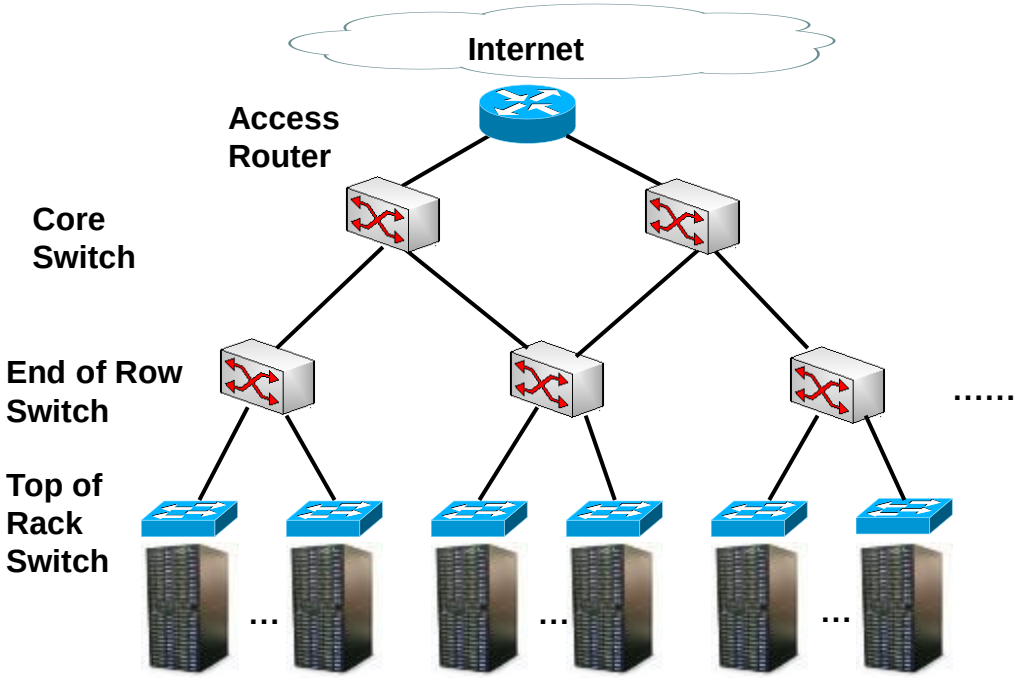
  Columbia University
- Jaime Cardenas, Carl Poitras, Brian Stern, Lawrence Tzuang, Michal Lipson

  Cornell University

Why silicon photonic networks?



Data Center Challenge



Traditional Data Center with Tree Architecture

- Bandwidth bottleneck
  - Traditional tree structure – aggregation bottleneck
  - New Data Center applications – increased “East / West” traffic
- Scaling current architecture “as is” is not economically feasible
- Power efficiency must be improved

Exascale

- Improving energy and power efficiency is the most formidable challenge
  - Exascale machines must be approximately 1000 times more energy efficient than current machines
- Recommended approaches by US National labs include
  - **Silicon photonic interconnects** –focus
  - **Hardware Software co-design** –focus

S. Hemmert et al. Exascale Hardware Architectures Working Group, NNSA Workshop: From Petascale to Exascale: R&D Challenges for HPC Simulation Environments, March 21, 2011

J. Shalf et al., Rethinking Hardware Software Codesign for Exascale Systems, Computer , Nov. 2011

Photonics is getting there

2012: Blue Gene/Q

IBM

Sequoia - (96) IBM Blue Gene/Q Racks
20.013 Pflops Peak … 1.572M Compute Cores … ~2026 MFlops/Watt

13

~8MW

- Optical transceivers tightly integrated, mounted within drawer
- 8 Hub/switch modules (8 x 56 optical modules)

Acknowledgment: A. Benner

F. Doany, Hot Interconnects 2012

Advances in Network Integration
Hybrid Networks

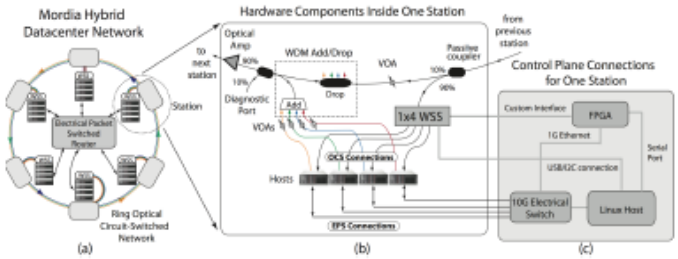
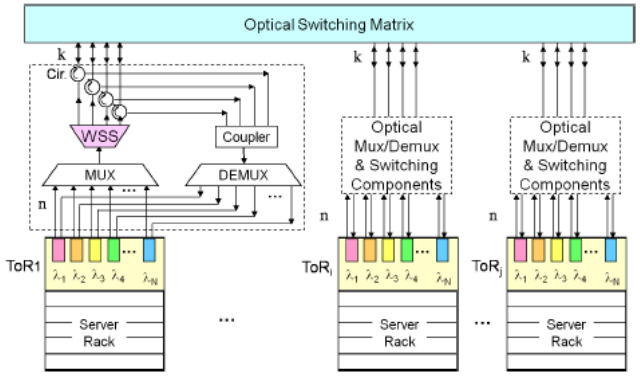
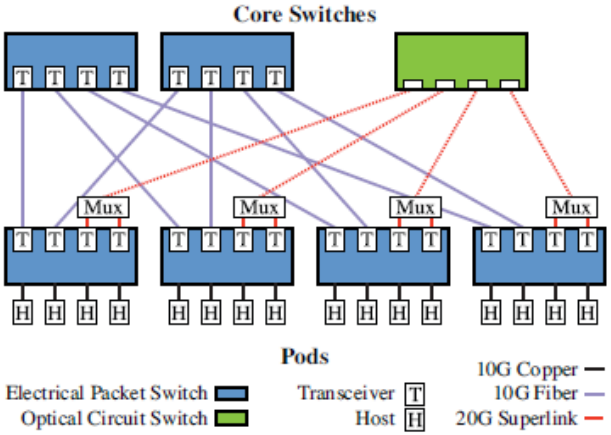


Fig. 1. System-level diagram of the Mordia network.

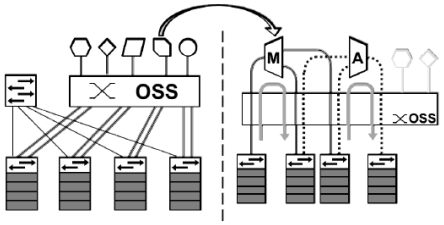
Farrington et al.
OFC 2103



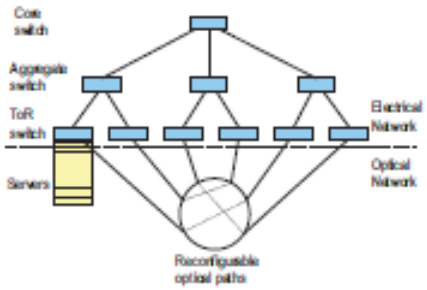
Chen et al. NSDI 2012



Farrington et al.
Sigcomm 2011



Wang et al.
CCR 2103



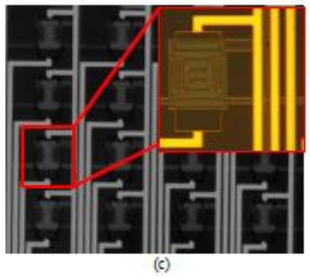
Wang et al. Hotnets 2009

Optical switch challenges

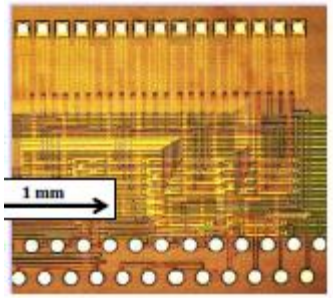
- Want a switch with
  - High bandwidth capacity
  - High port count
  - Low power
  - Small footprint
  - Fast switching speed
  - Low cost

Silicon Photonics

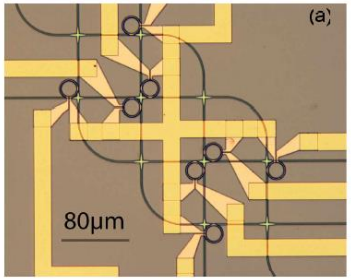
- CMOS compatible **silicon photonics**
  - Opens new possibilities for multiwavelength (WDM) photonic circuits
  - Integration for low power consumption, high bandwidth, overcoming pin count limitations
  - Area of intense research, IBM, MIT, Intel, TUE, UCSB, Oracle, Mellanox and many others



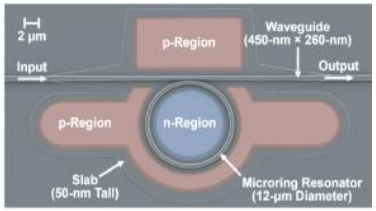
DasMahapatra et al.
OFC 2013



Lee et al. OFC 2013

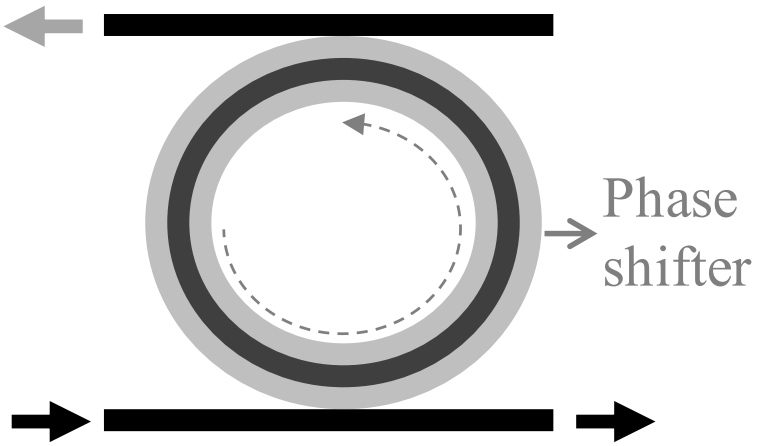


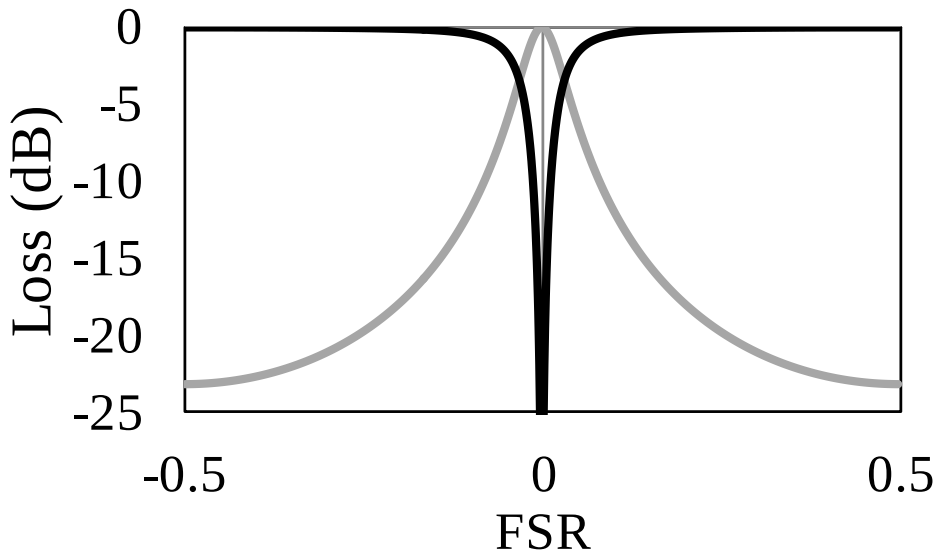
Droz et al.
Opt. Express 2008



Padmaraju et al.
Opt. Express 2012

Single ring resonator





Refractive index tuning by

- **Carrier injection** – *fast but low tuning range*
- **Thermo-optic** – *limited speed but full tuning*

Challenge

- Temperature and Fabrication dependent resonant wavelength

Switching element

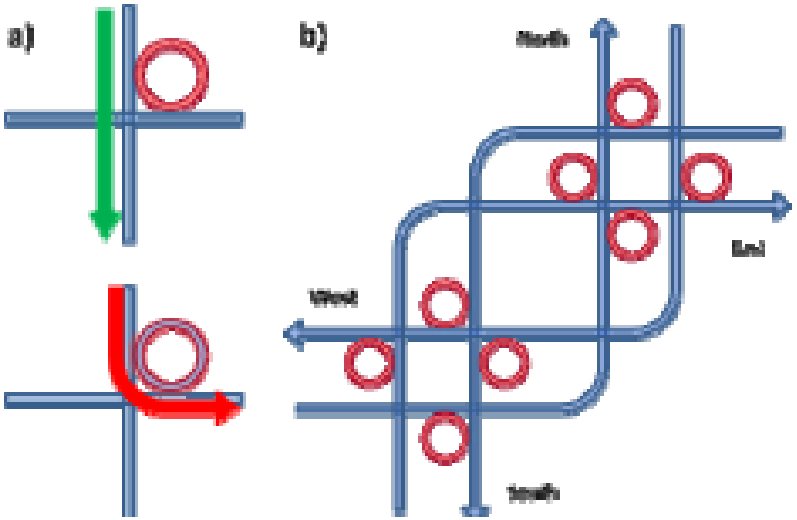
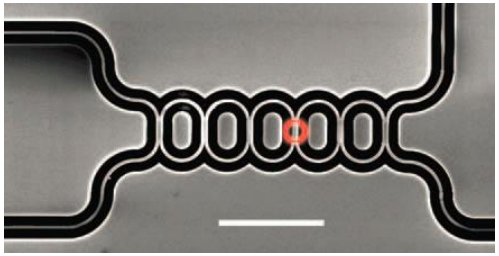
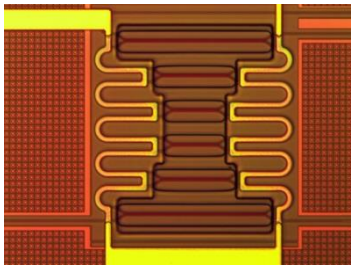


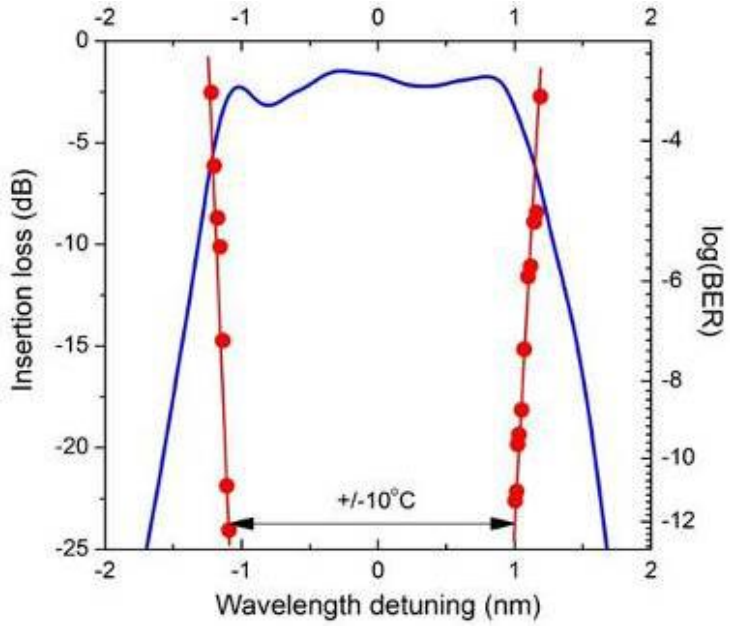
Fig. 1: (a) Atomic switching element depicting message propagation in two possible states. (b) 4×4 non-blocking nanophotonic switching node.

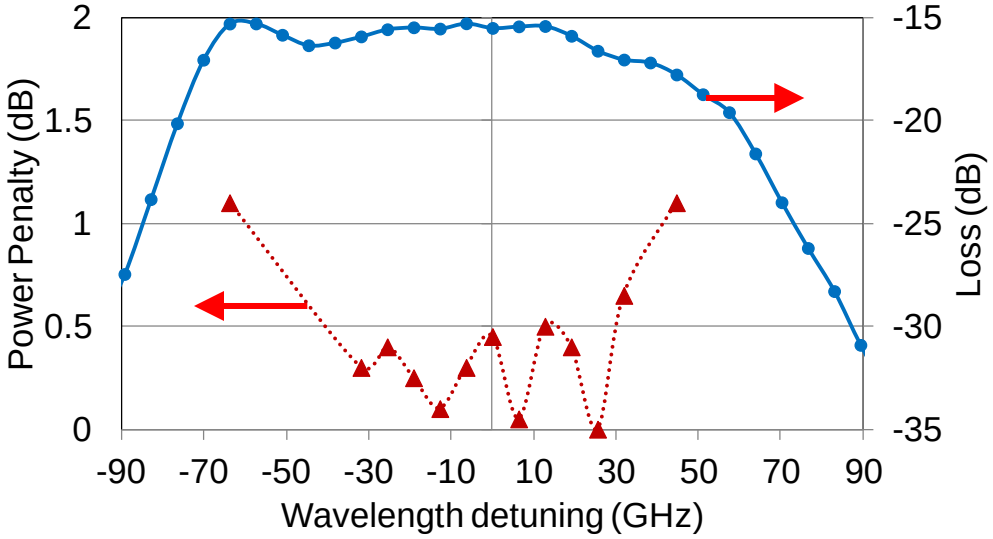
Wang et al. OFC 2008

Higher order resonators: Resilience to temperature and wavelength detuning







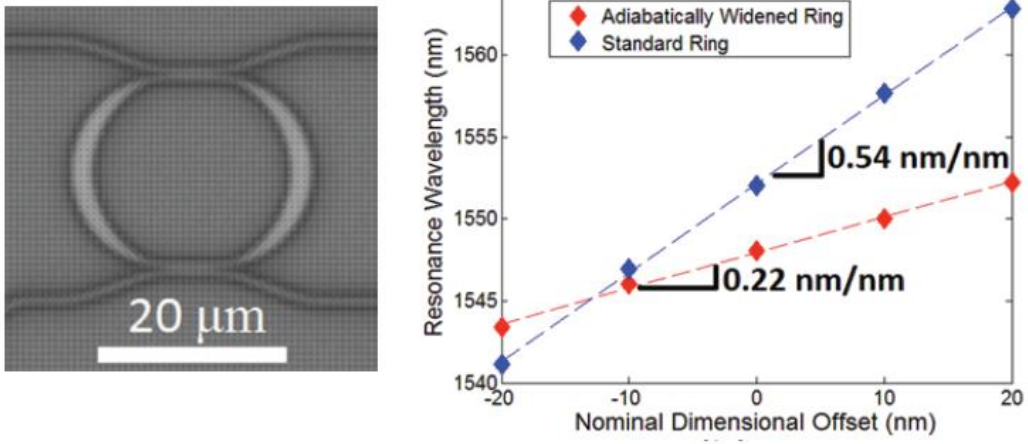


- Total detuning range of greater than 100 GHz is possible with penalty <1.1 dB

Ref: Vlasov (IBM)., Tutorial ECOC, 2008

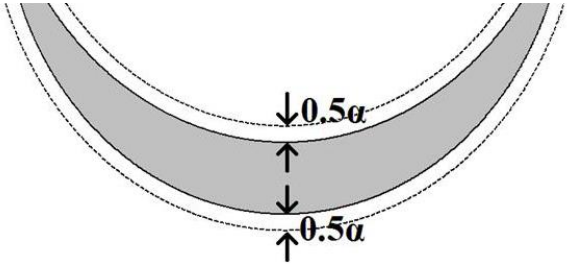
Ref: A.Rohit et al., PTL, Dec 2012

Minimizing Fabrication-induced mismatch



(a) (b)

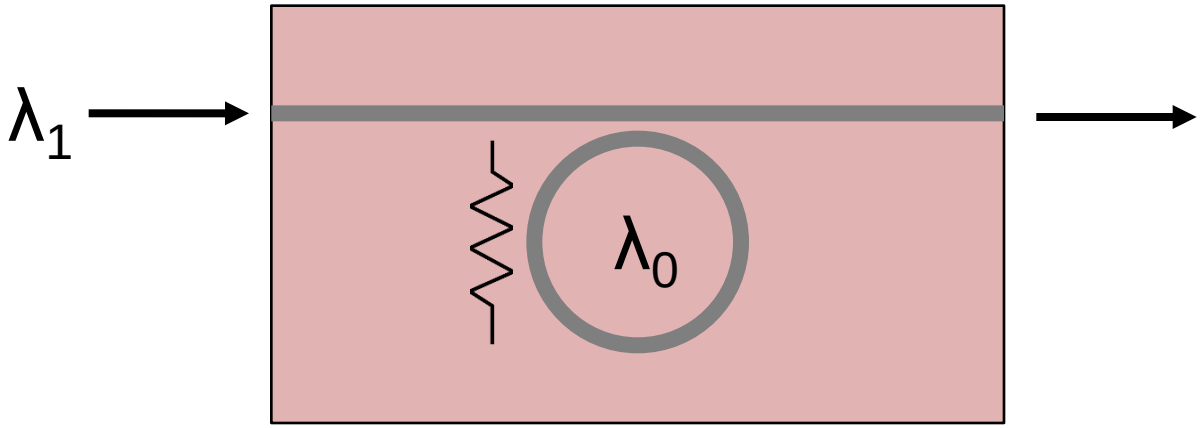
Fig. 1. (a) Optical microscope image of a fabricated adiabatically widened microring in SOI with an oxide cladding. (b) Resonance wavelength shift as a function of the nominal dimensional offset. (From [6])

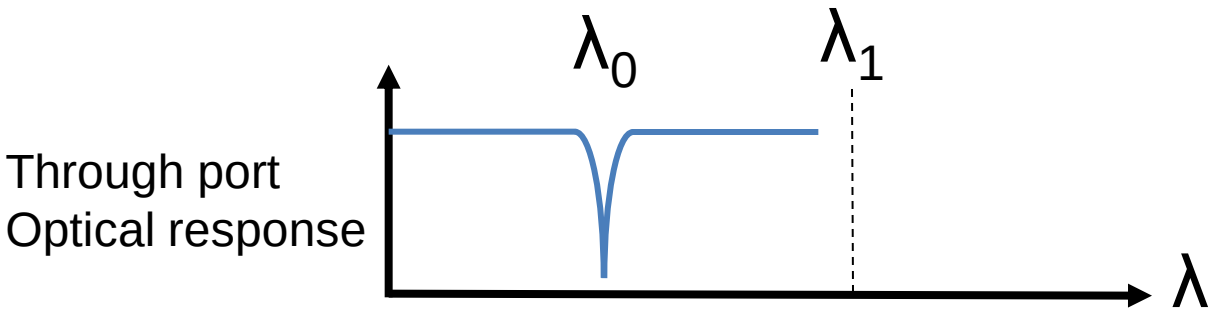


α = Dimensional offset

Ref: Poon et al., Redesigning active and passive microring resonators

Initialize microring wavelength using heaters

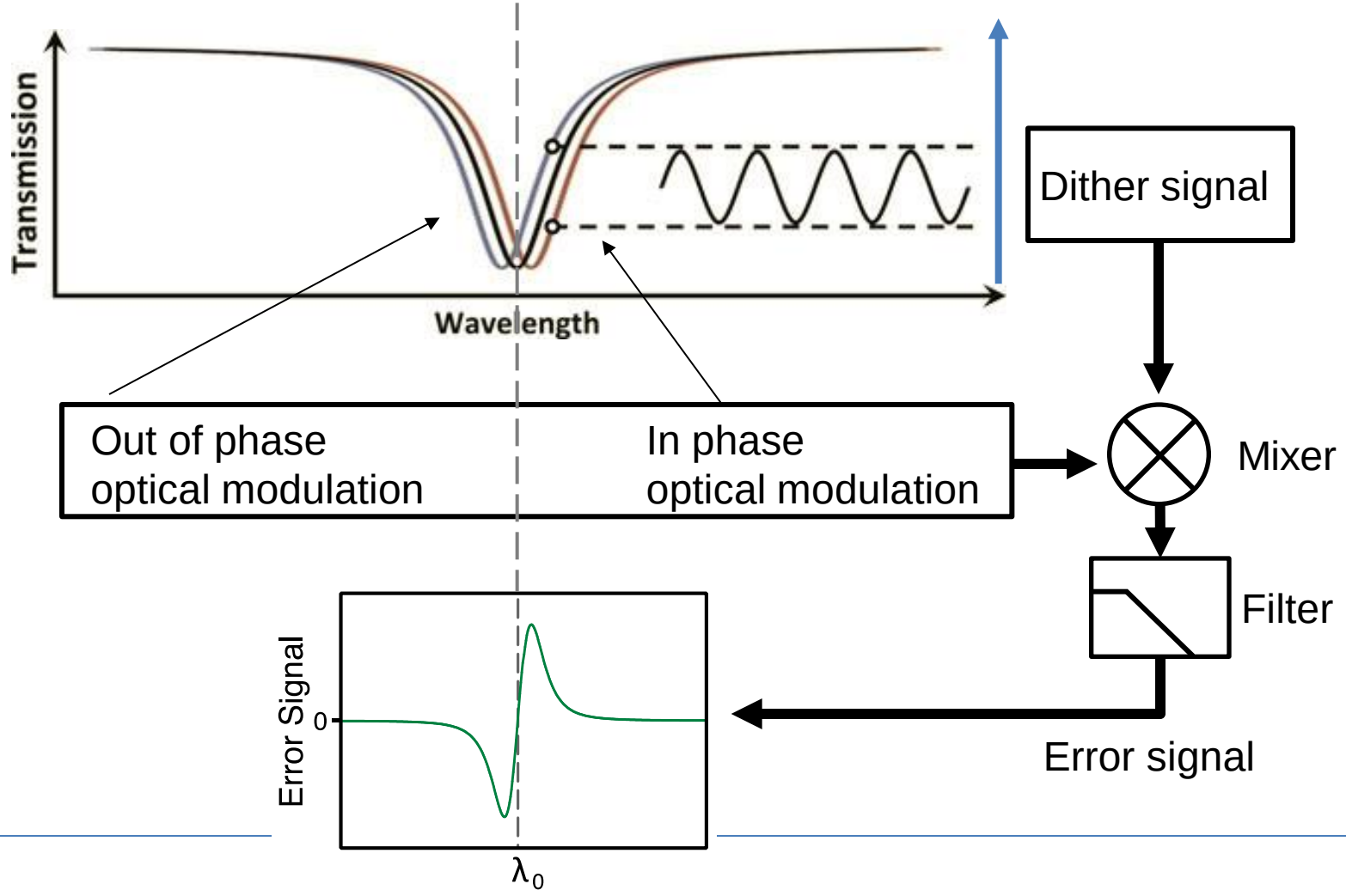




[A. Krishnamoorthy 2011]

[M. Watts 2010]

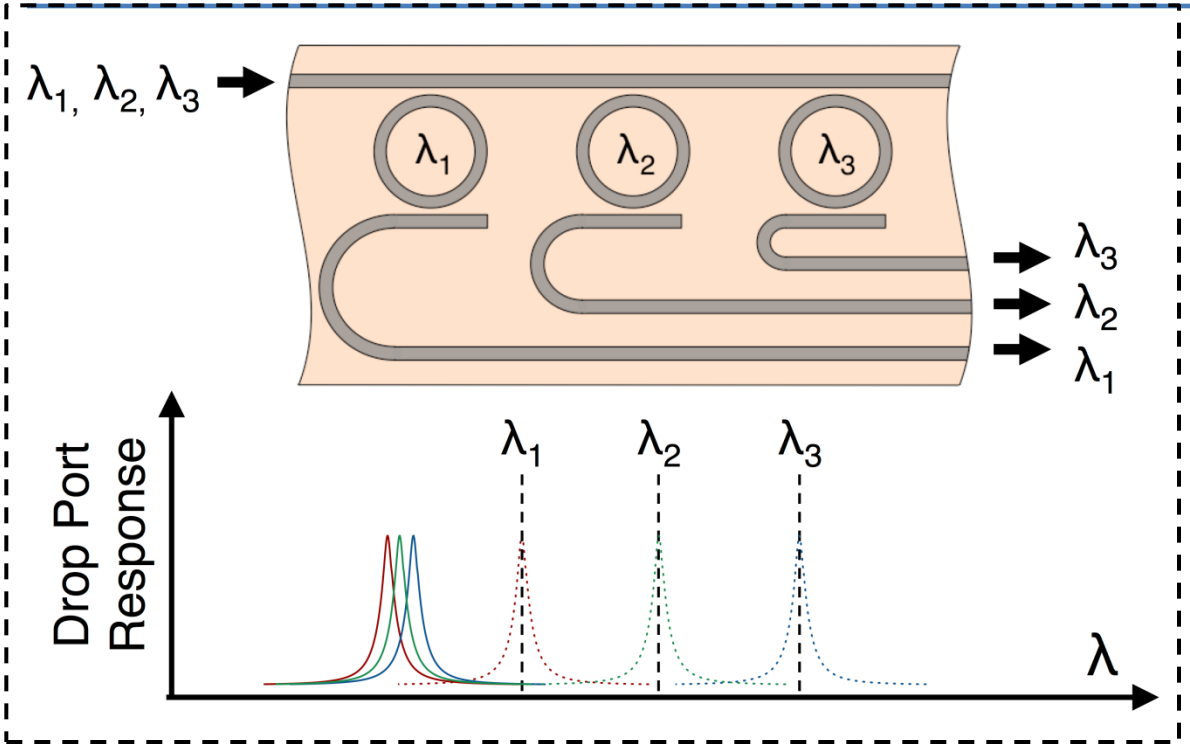
Dither based locking



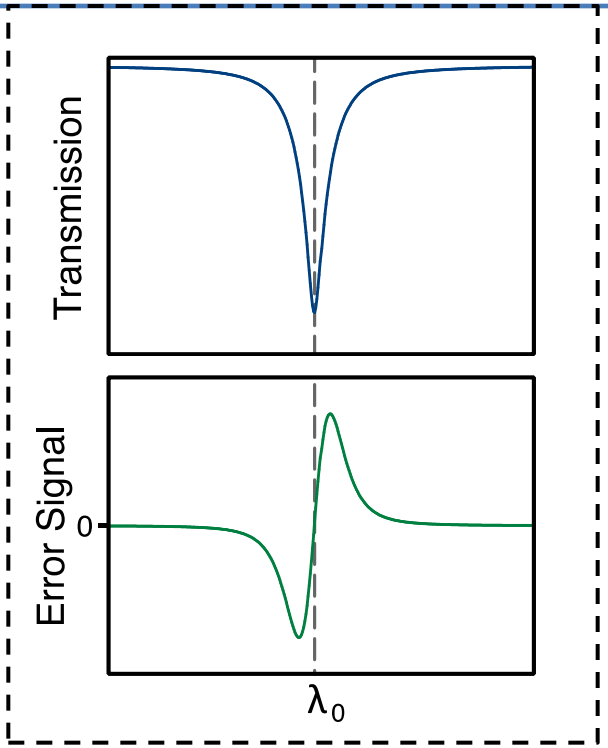
[K. Padmaraju 2013]

Wavelength Locking of a WDM Microring Demux

WDM Microring Demultiplexing Filter



Generation of Error Signals

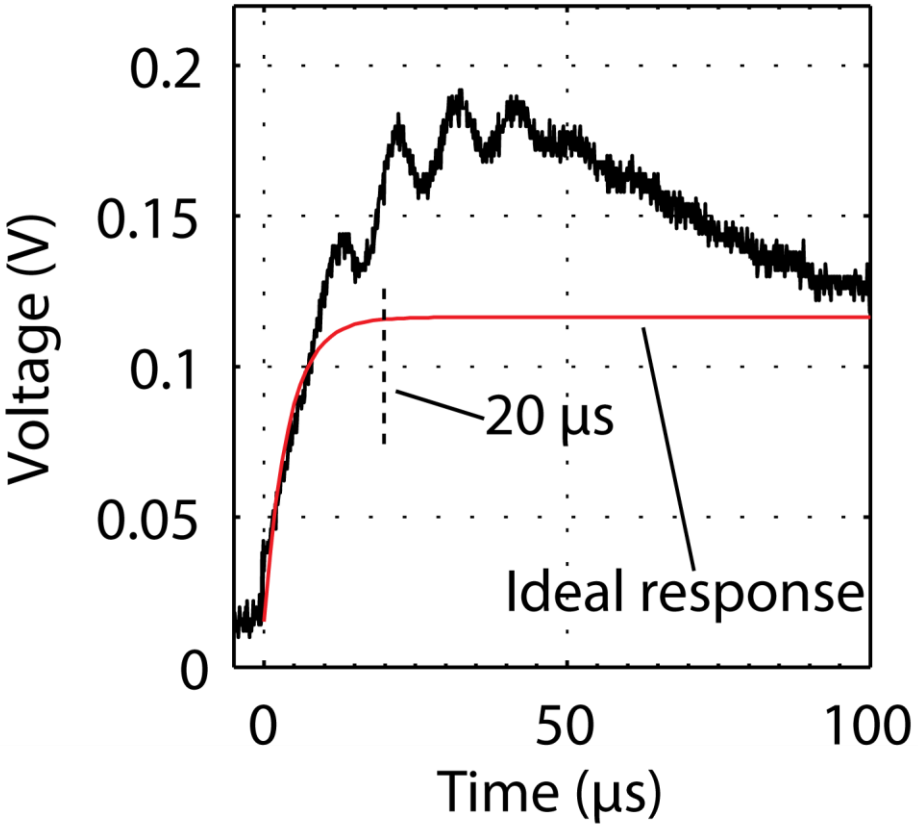


Experimental Demonstration: Initialization of a Microring Demultiplexer

- Integrated heaters must tune microring filters to WDM channels
- We have developed techniques (in mixed-signal circuitry) to wavelength lock & thermally stabilize microring resonators.
- Required migration of our techniques to an FPGA platform such that it can be incorporated into the FPGA testbed.

K. Padmaraju et al., OFC 2014

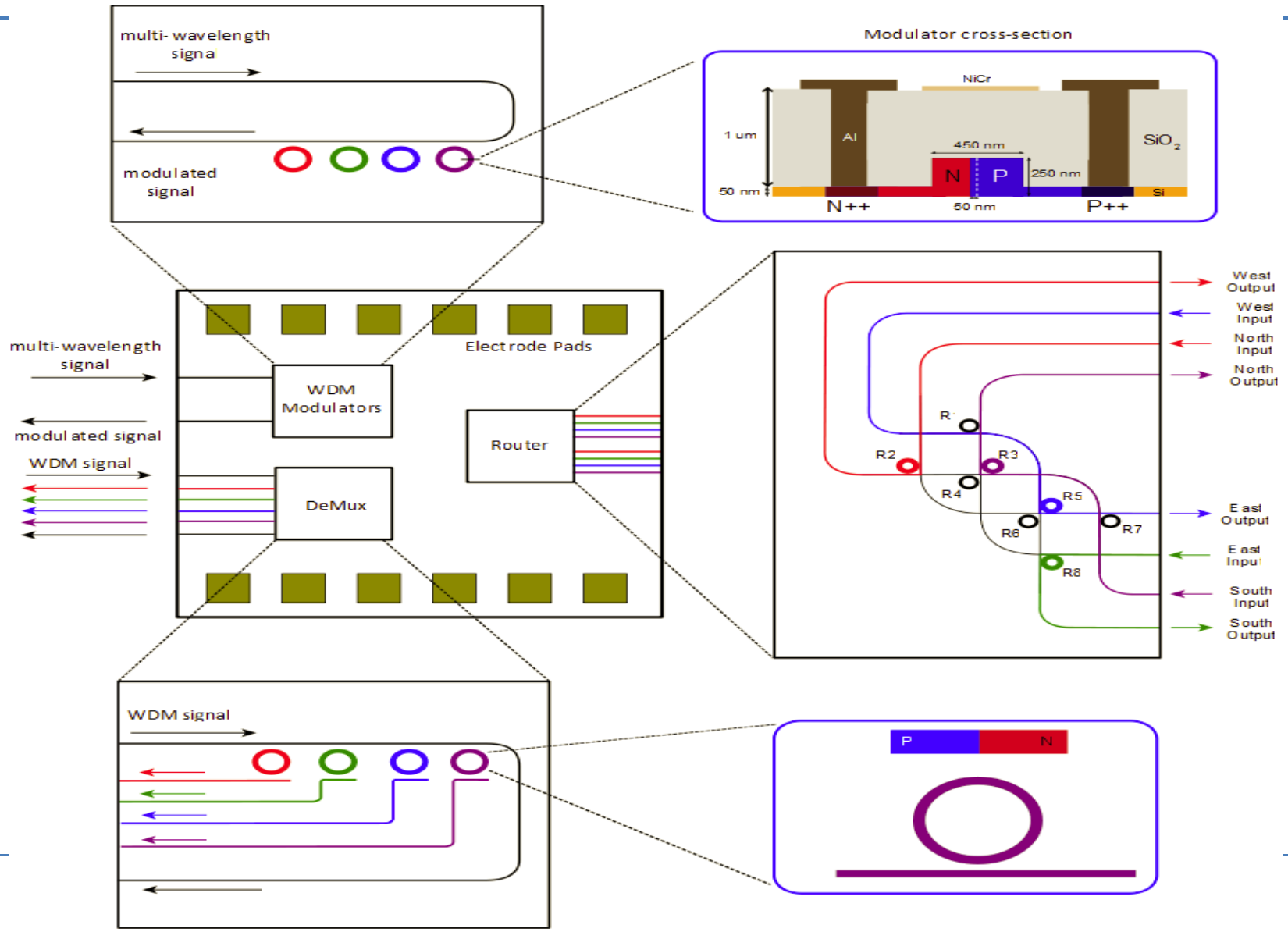
Feedback controller settling time



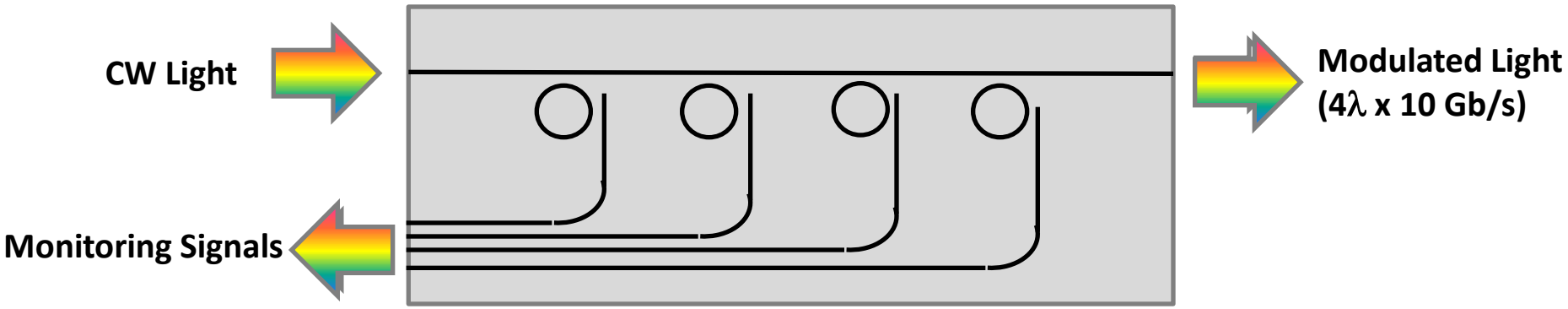
- Feedback lock is order magnitude slower than ideal
- Limited by error signal generation, and the dither frequency

Zhu et al. Optical Interconnects 2014

FPGA Programmable SiP **Interconnection Network**

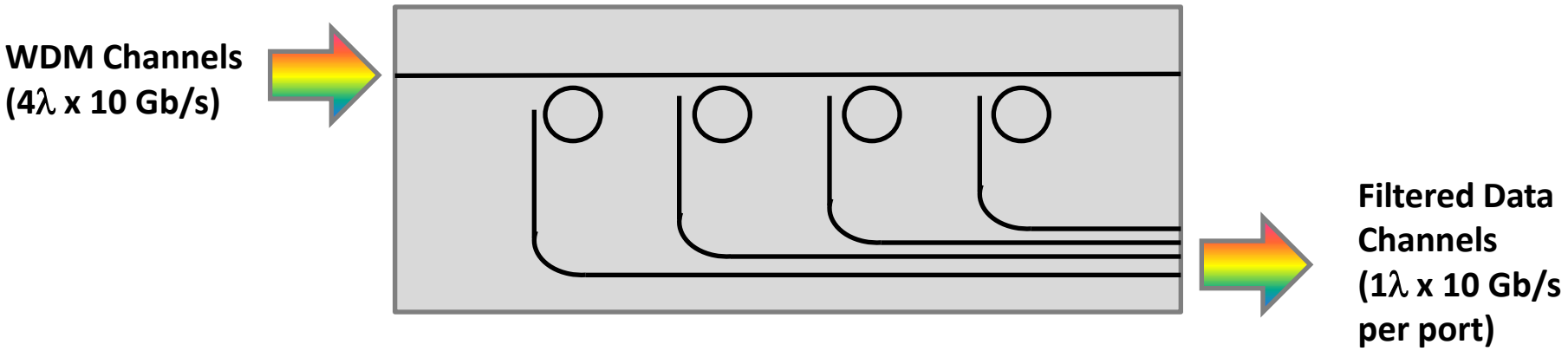


Microring Modulator Array



- Actively thermally stabilized variant:
- > 5 dB ER with ~ 2 Vpp drive
- < 5 dB IL (excluding fiber coupling)
- To include heaters (local and global)
- To include on-chip photodetectors on drop ports

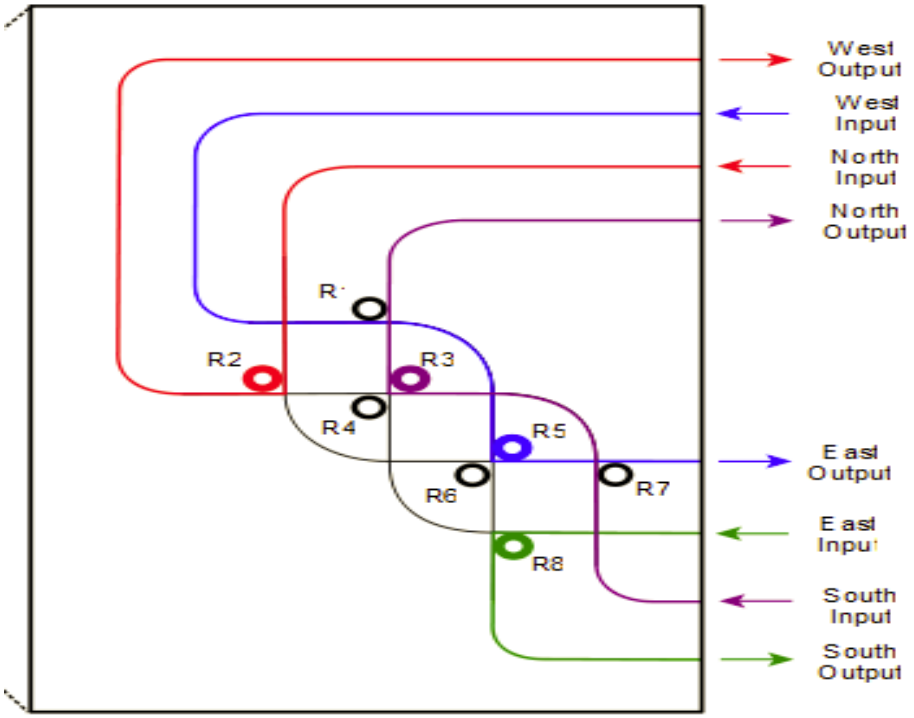
WDM Demux Array



- Separates each channel to a different output port for detection by off-chip photodetectors
- Gen 2 includes additional on-chip slow photodetectors for thermal stabilization
- Low penalty for filtering 10-Gb/s channels
- < 1 dB IL
- < -20 dB crosstalk between wavelength channels

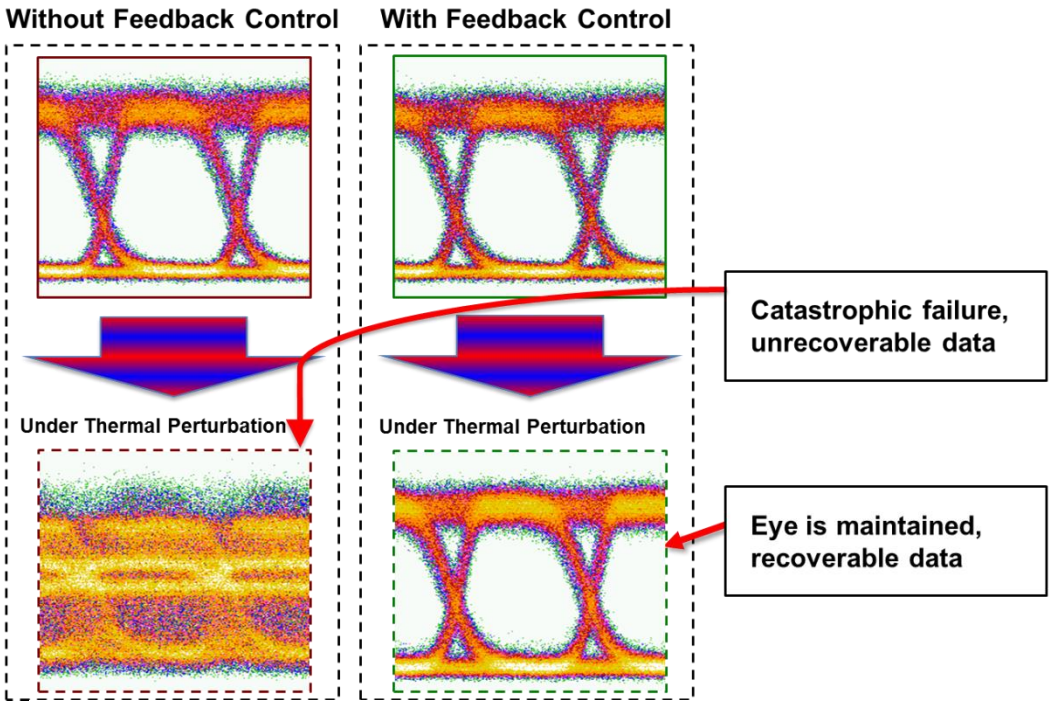
Switch

- 4x4 switch (non-blocking) for routing between four inputs and four outputs
- 8 rings with thermal tuning (resonance tuning) and MZI's (coupling tuning)
- 4 wavelength channels near λ = 1550 nm, FSR = 3 nm (channel spacing)

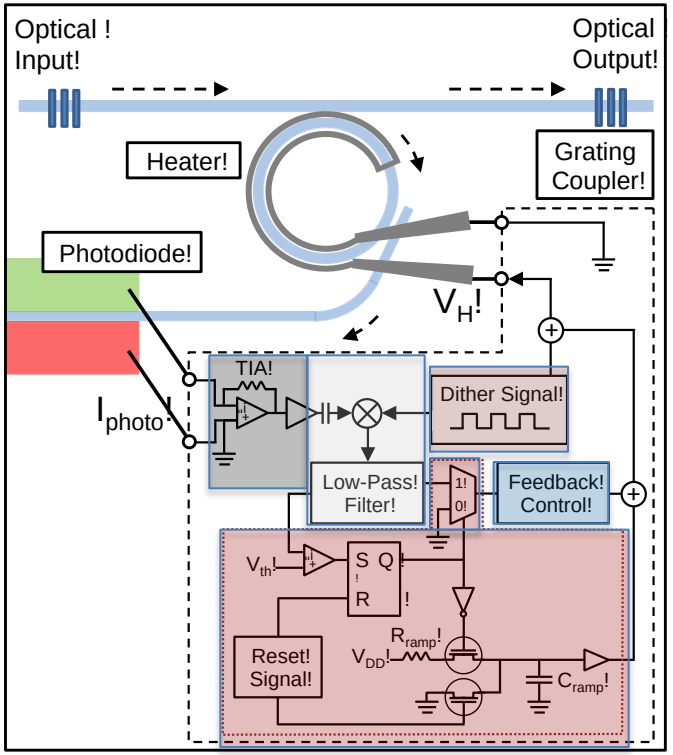


Physical Level Control for SiPh

- Photonic devices require dedicated driving and stabilization.
- Digital logic for
  - Balanced coding / decoding
  - Thermal stabilization
  - Switch dedicated driving signals

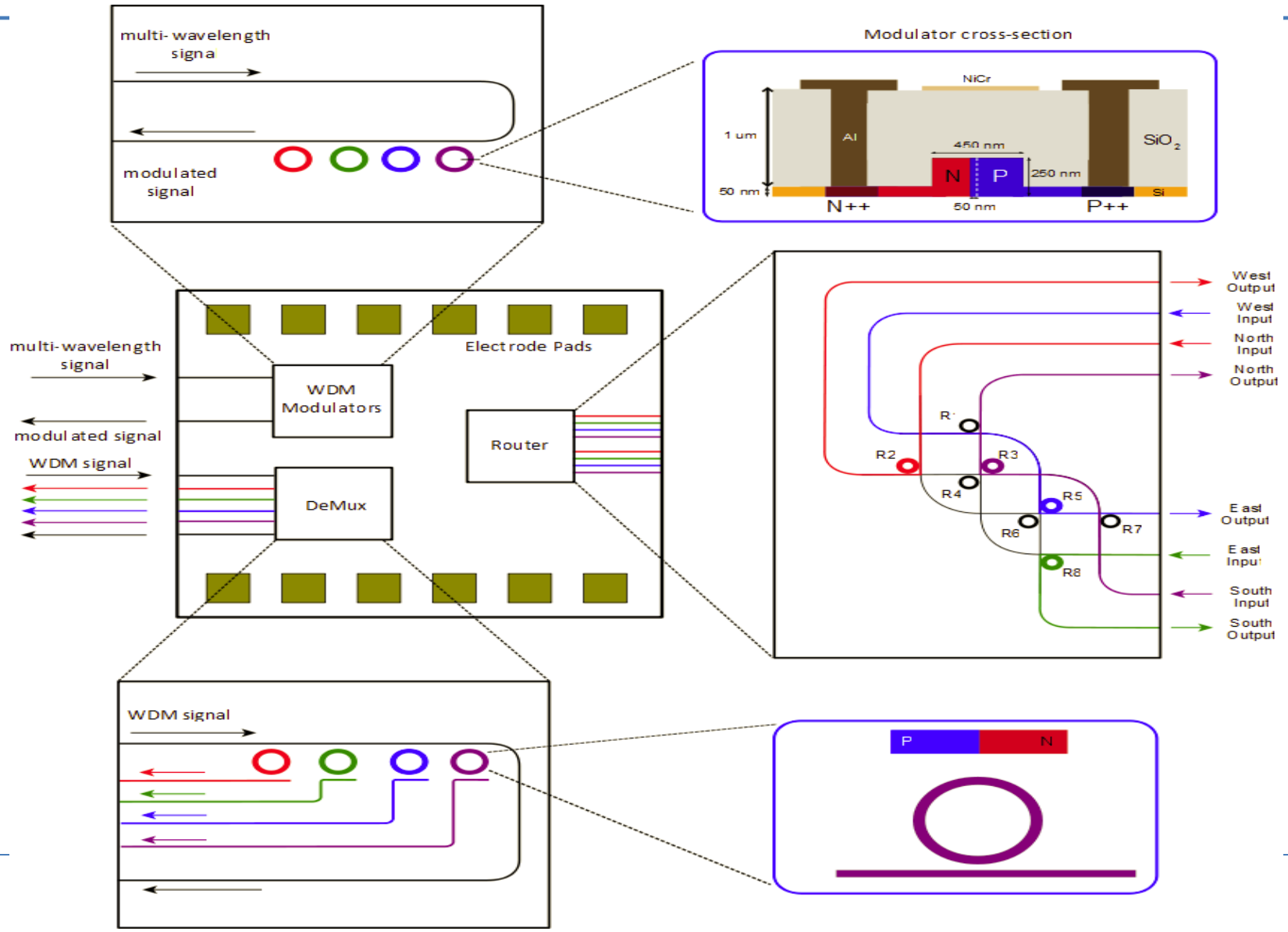


Thermal stabilization through optical feedback



ADC input
DSP functionality
DAC / PWM output
Digital PID controller
Standard logic elements

FPGA Programmable SiP **Interconnection Network**



Mach Zehnder based switch

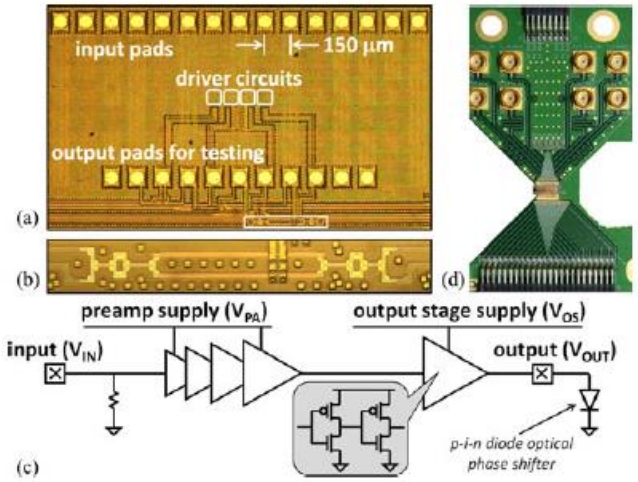


Fig. 1. (a) Die image showing a four-channel driver test site. (b) Magnified image of the WIMZ switch outlined in (a). (c) Schematic of the digital switch driver. (d) Custom test card.

TABLE II
CMOS-DRIVEN 4 × 4 MZ FABRIC EXTRACTED POWER AND ENERGY PERFORMANCE

Supply Rail	Static Power[1] (mW)	Switching Energy[2] (pJ/cycle)
Logic	< 2.1	---
PA	2.4	82
OS	7.6	24
Thermo-Optics	34.4	---
Total	46.5	106

[1]Average power assuming a 50% state probability.
[2]Includes turn-on and turn-off dissipations.

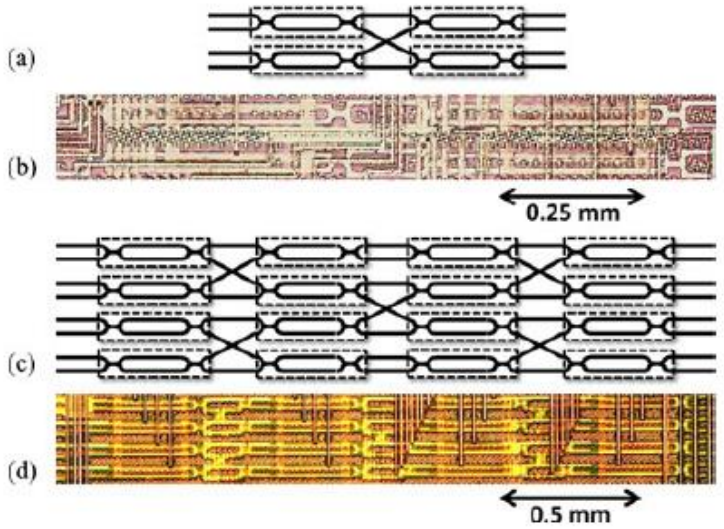


Fig. 6. (a) Topological arrangement and (b) micrograph of the 4 × 4 fabric. (c) Topological arrangement and (d) micrograph of the 8 × 8 fabric. The dashed lines in (a) and (c) represent 2 × 2 MZ-based switching elements.

B. Lee et al. Journal Lightwave Technology 2014

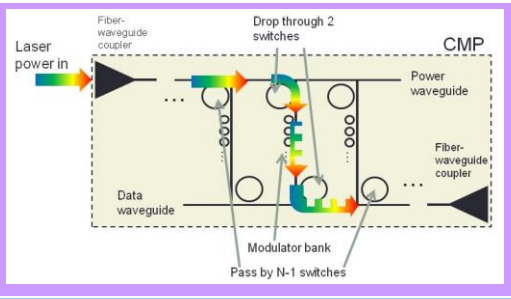
Simulations

- Scaling
  - Most experiments done on smaller scale testbeds
  - Working data centers aren't turned off to do an experiment.
- Effects of changes without complete implementation
  - Current application performance on new architecture
  - Studies at extreme scales and evaluation of new programming models and algorithms

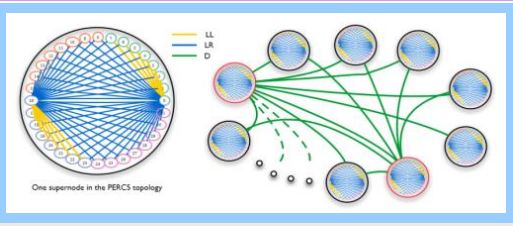
Photonic-Enabled Systems: Multi-Level Co-Design

PhoenixSim: Design, Modeling and Simulation Environment

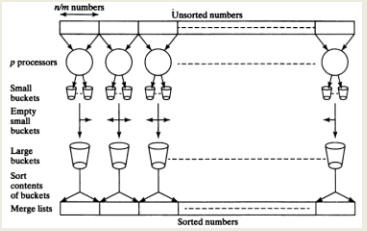
- **Physical link layer:**
  - SiP components modeling
  - Link bandwidth maximization
  - Optical power budget validation



- **Network layer**
  - Optical data flow, switching, routing protocols
  - Network performance analysis



- **Application layer**
  - BW and data flow application mapping
  - Optically enabled algorithm re-design
  - Large scale application simulation



S. Rumley et al., ICTON 2013, Optical Interconnects 2014

Silicon Photonics Challenges

- Photonics has demonstrated enormous bandwidth capabilities for telecoms
- But must reduce size, cost, power for datacomms
- Current challenges for Silicon Photonics
  - Microring stability and resilience to fabrication variations
  - System integration
    - Building interfaces to computer networks
    - Demonstrating performance improvements for users

Thank you