Designing Future VLSI Systems with Monolithically Integrated Silicon-Photonics

Vladimir Stojanović
University of California, Berkeley

SSCS DL Lecture
University of Texas, Austin
November, 2013

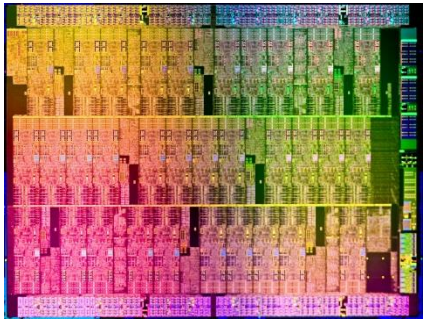
Acknowledgments

- Milos Popović (Boulder), Rajeev Ram, Michael Watts, Hanqing Li (MIT), Krste Asanović (UC Berkeley)
- Jason Orcutt, Jeffrey Shainline, Christopher Batten, Ajay Joshi, Anatoly Khilo
- Mark Wade, Karan Mehta, Erman Timurdogan, Jie Sun, Cheryl Sorace, Josh Wang
- Michael Georgas, Jonathan Leu, Benjamin Moss, Chen Sun, Yu-Hsin Chen
- Yong-Jin Kwon, Scott Beamer, Yunsup Lee, Andrew Waterman, Miquel Planas
- Roy Meade, Gurtej Sandhu and Fab12 team (Zvi, Ofer, Daniel, Efi, Elad, …)

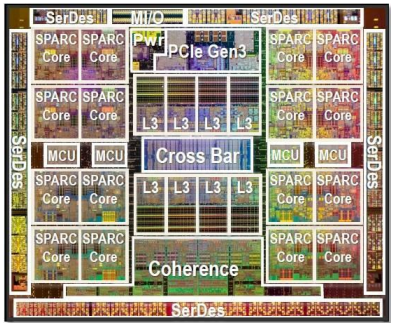
- DARPA, Micron, NSF and FCRP IFC
- IBM Trusted Foundry, CNSE Albany, Solid-State Circuits Society

Chip design is going through a change

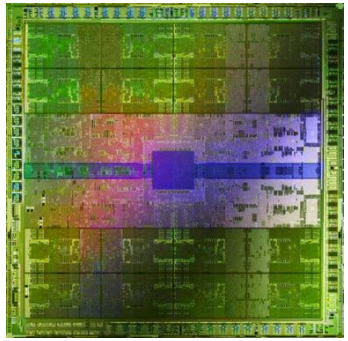
- Already have more devices than can use at once
- Limited by power density and bandwidth



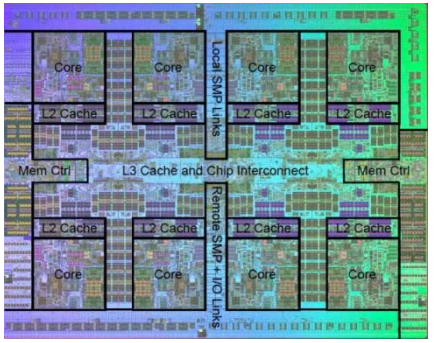
Intel Knights Corner
50 cores, 200 Threads



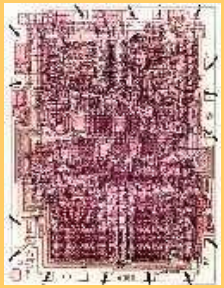
Oracle T5
16 cores, 128 Threads



Nvidia Fermi
540 CUDA cores



IBM Power 7
8 cores, 32 threads

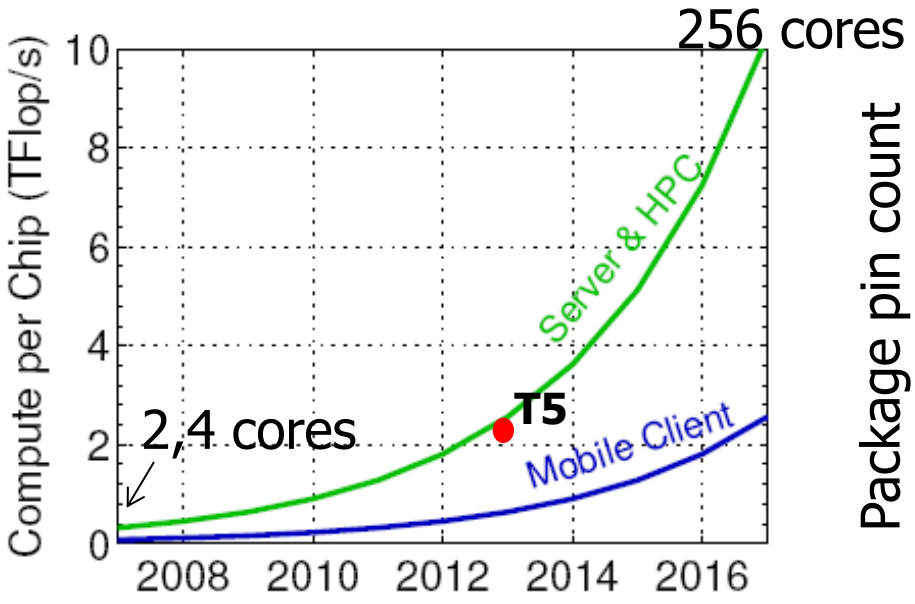


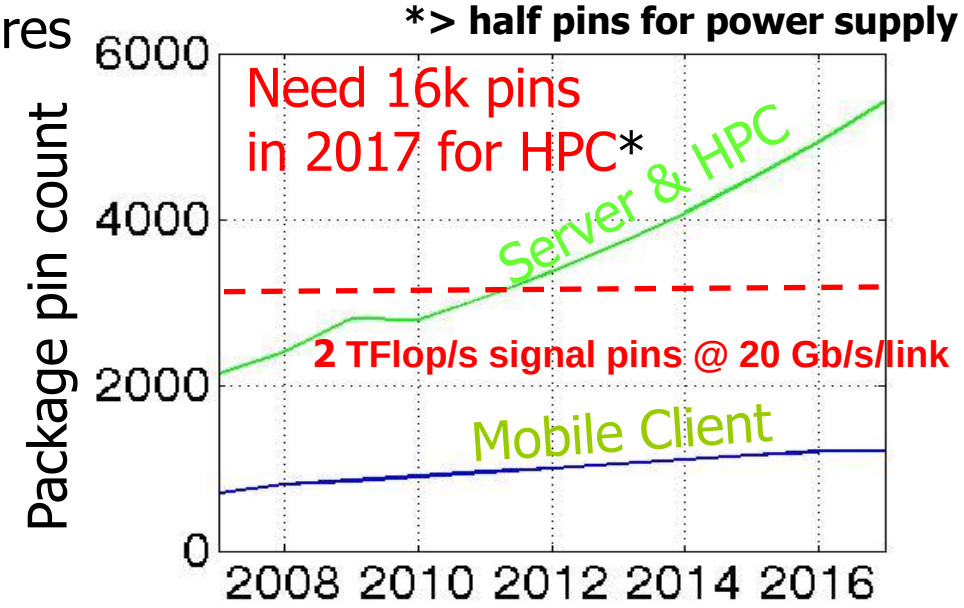
Intel 4004 (1971):
4-bit processor,
2312 transistors,
~100 KIPS,
10 micron PMOS,
11 mm^2 chip

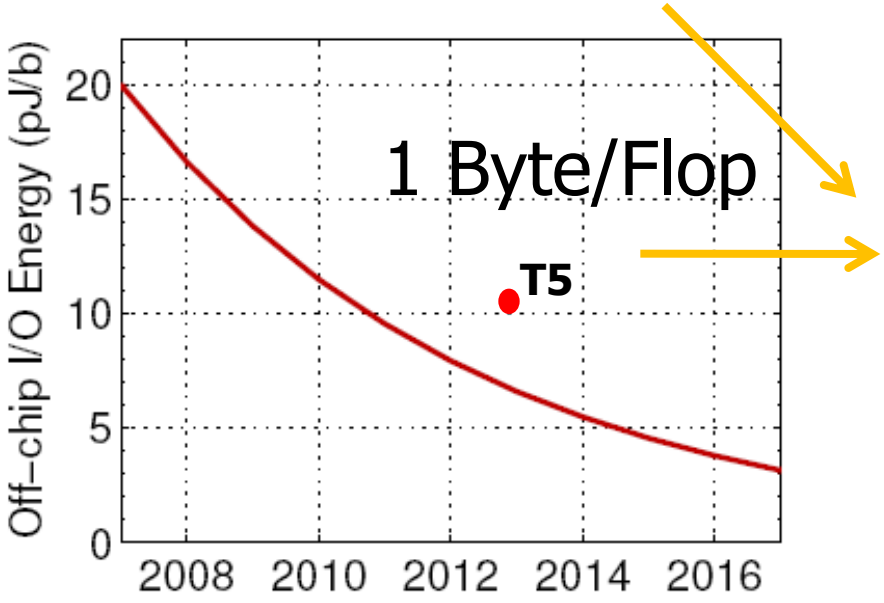
1000s of processor cores and accelerators per die

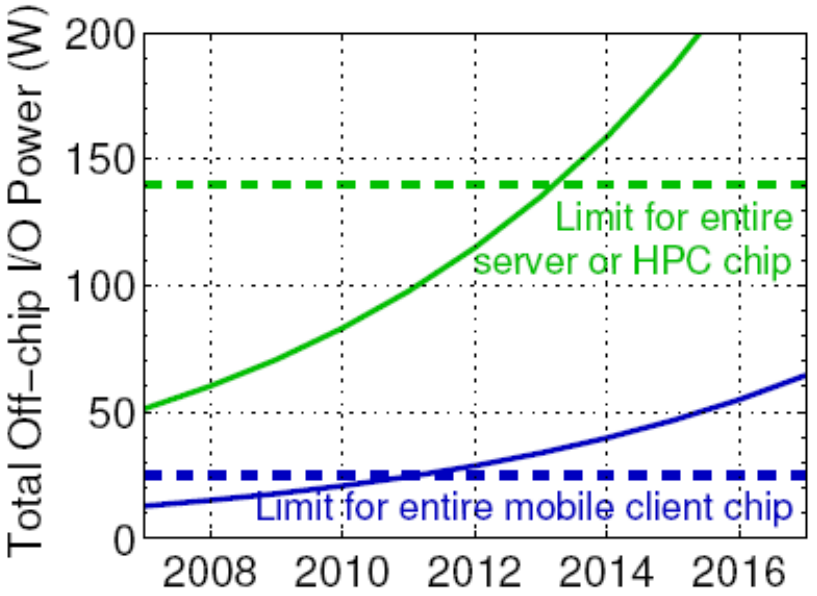
"The Processor is the new Transistor" [Rowen]

Bandwidth, pin count and power scaling

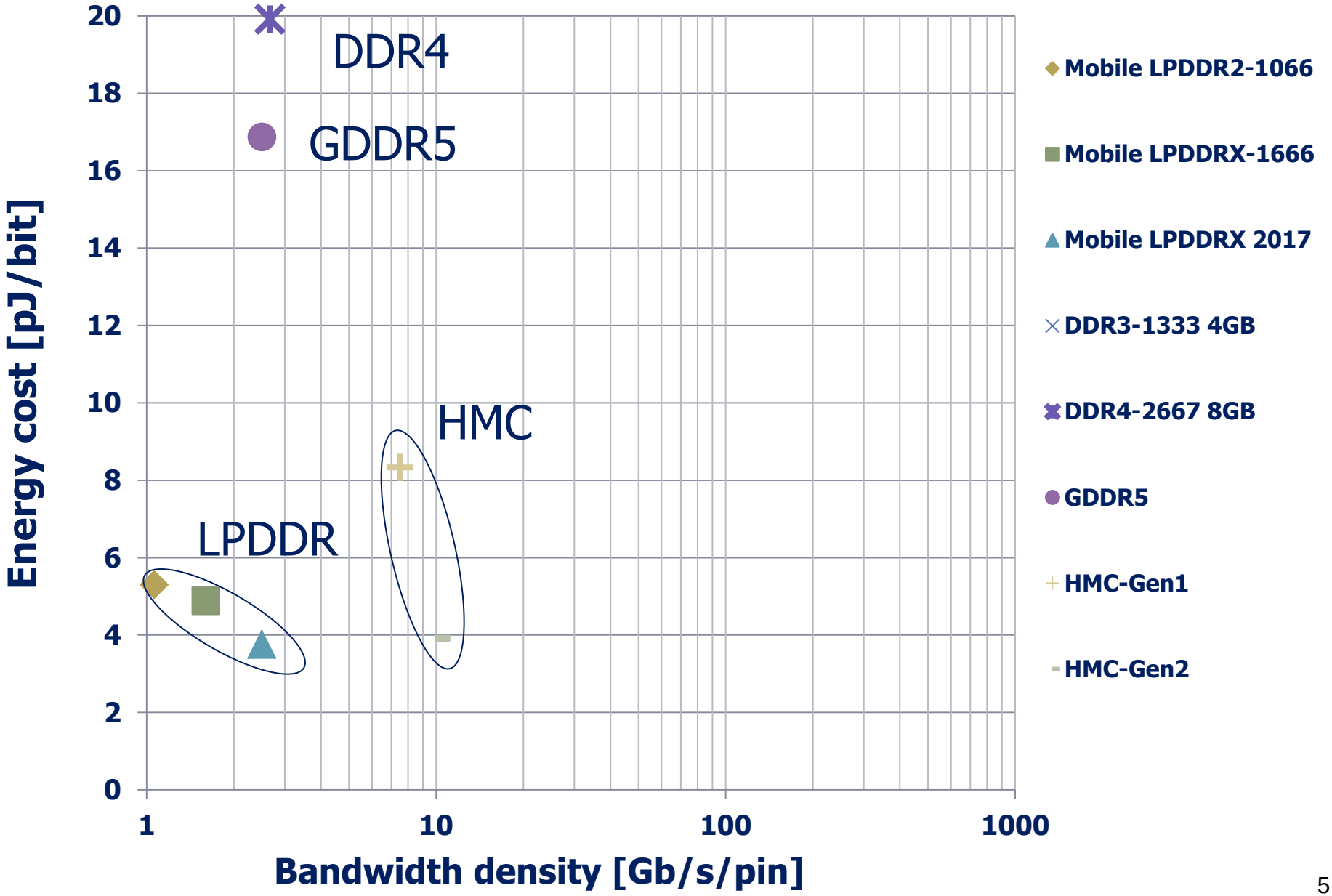




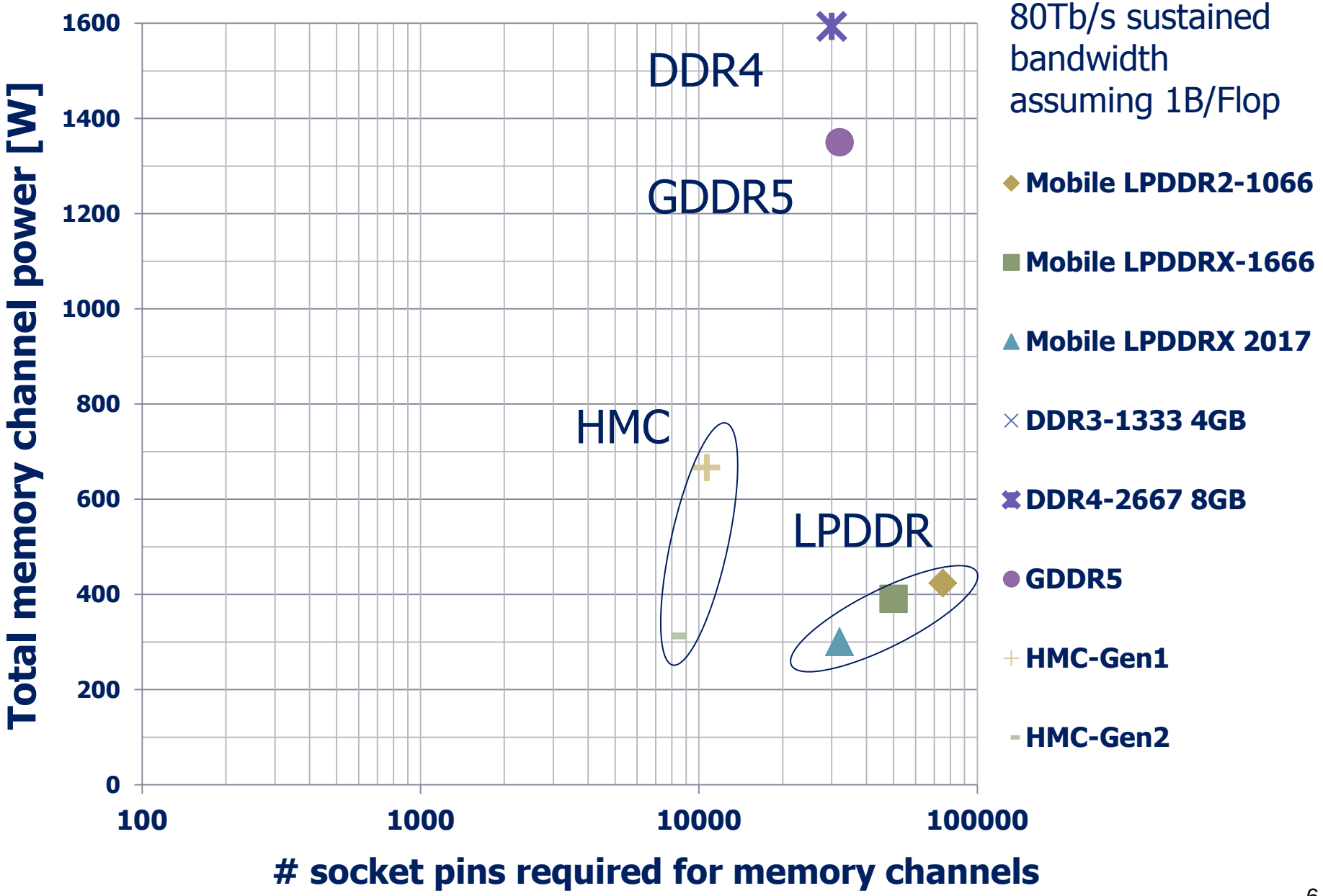




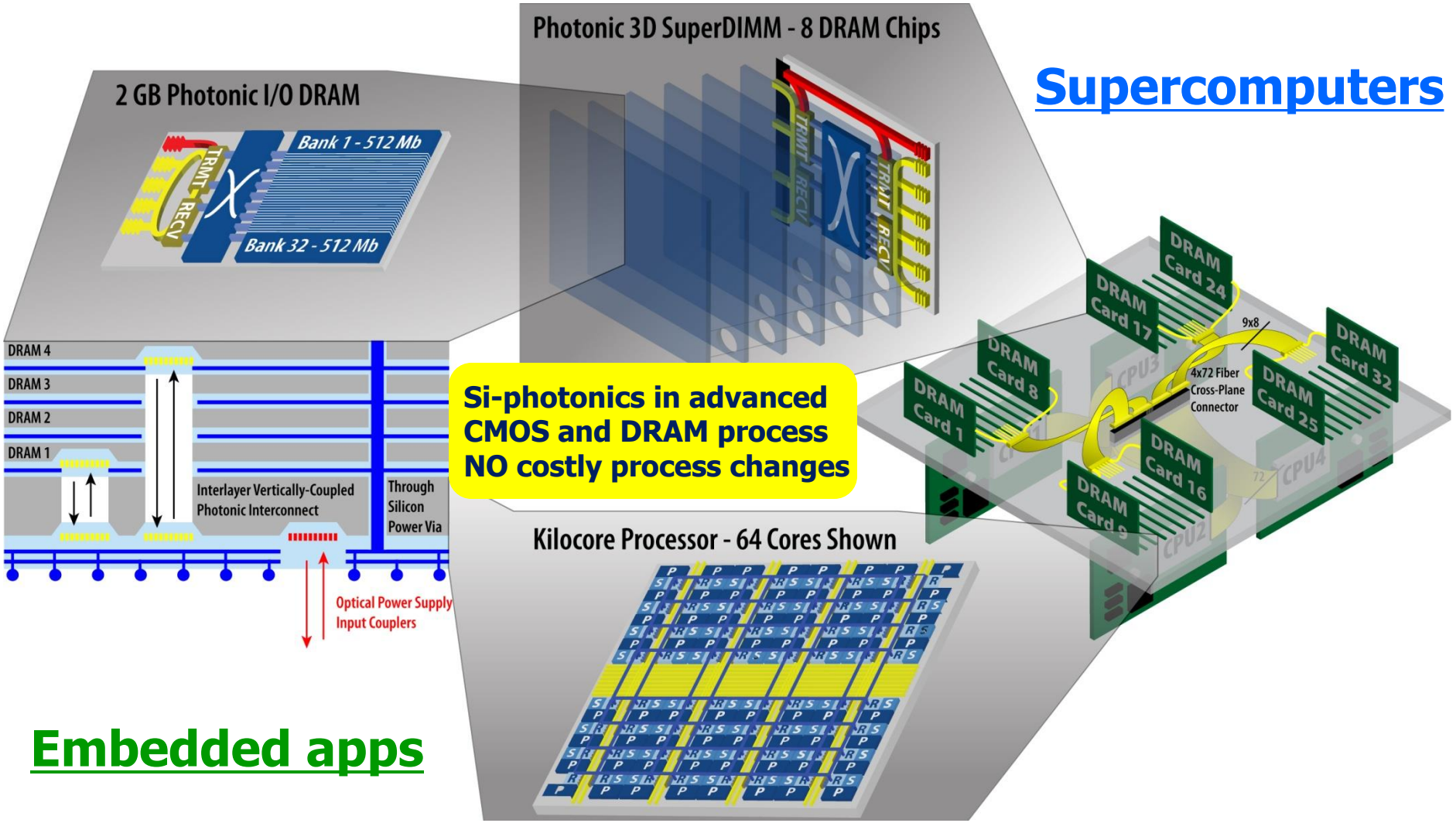
Memory interface scaling problems: Energy-cost and bandwidth density



Power and pins required for 10TFlop/s



Monolithic Si-Photonics for core-to-core and core-to-DRAM networks



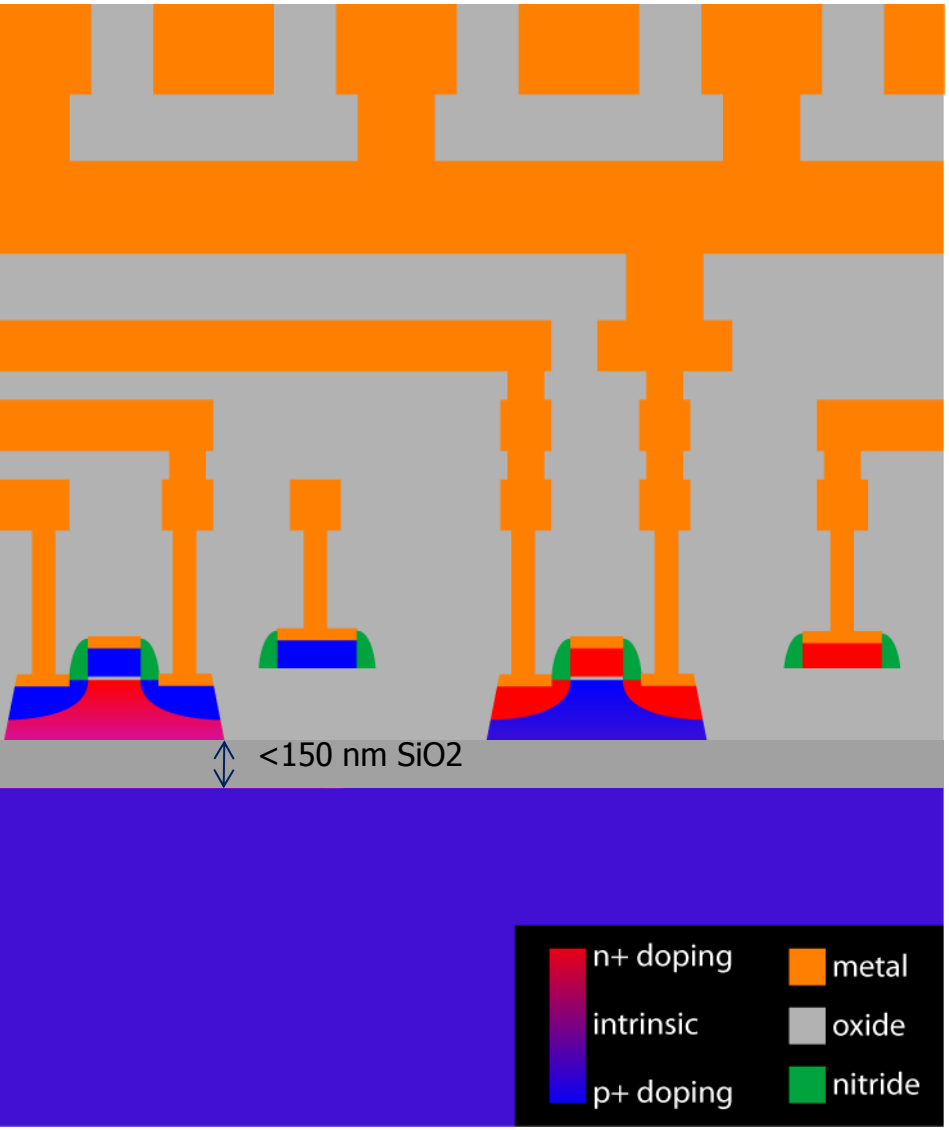
Supercomputers

Embedded apps

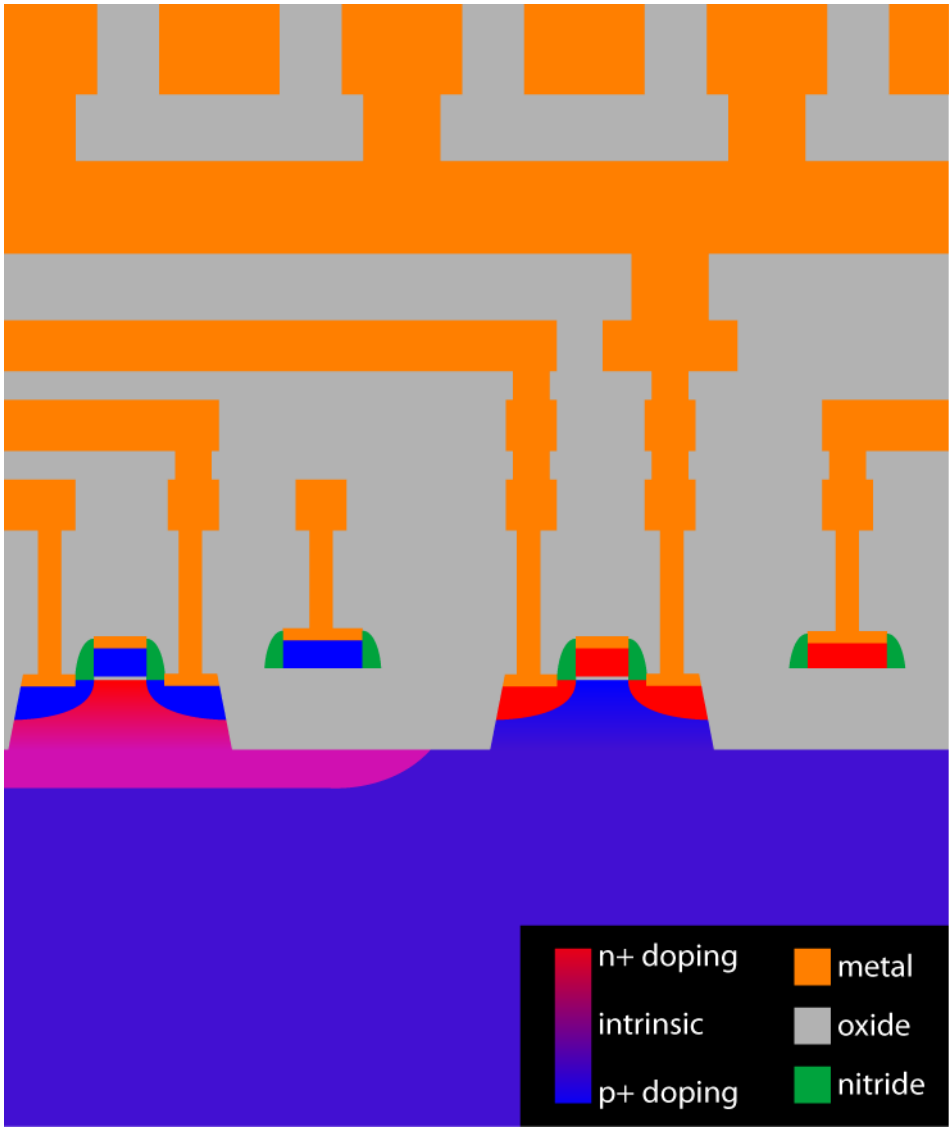
Bandwidth density – need dense WDM

Energy-efficiency – need monolithic integration

Monolithic CMOS photonic integration

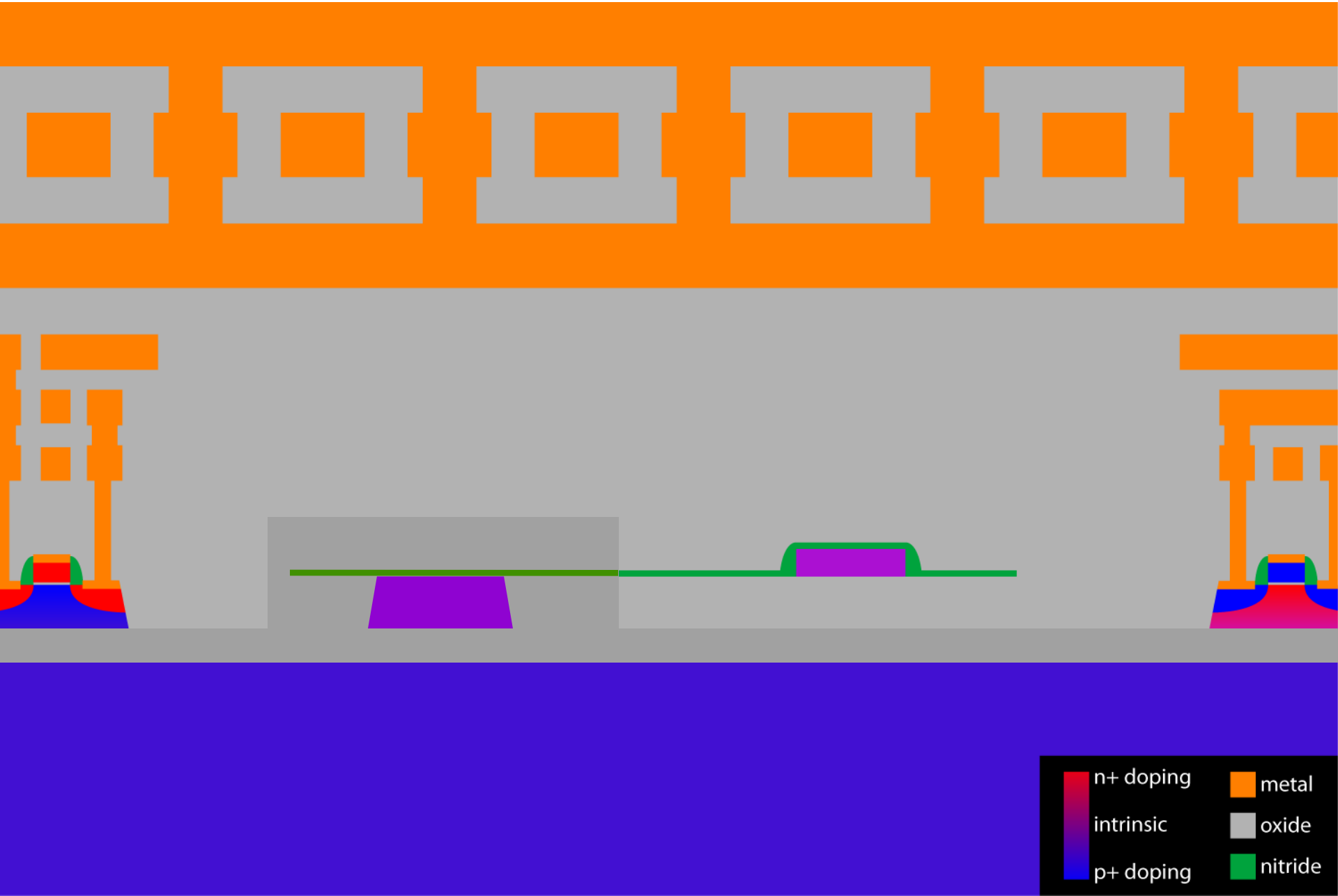


Thin BOX SOI CMOS Electronics

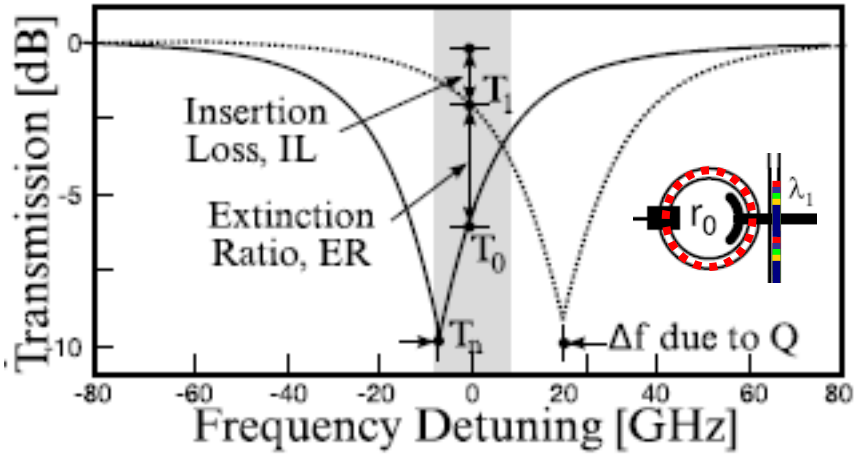


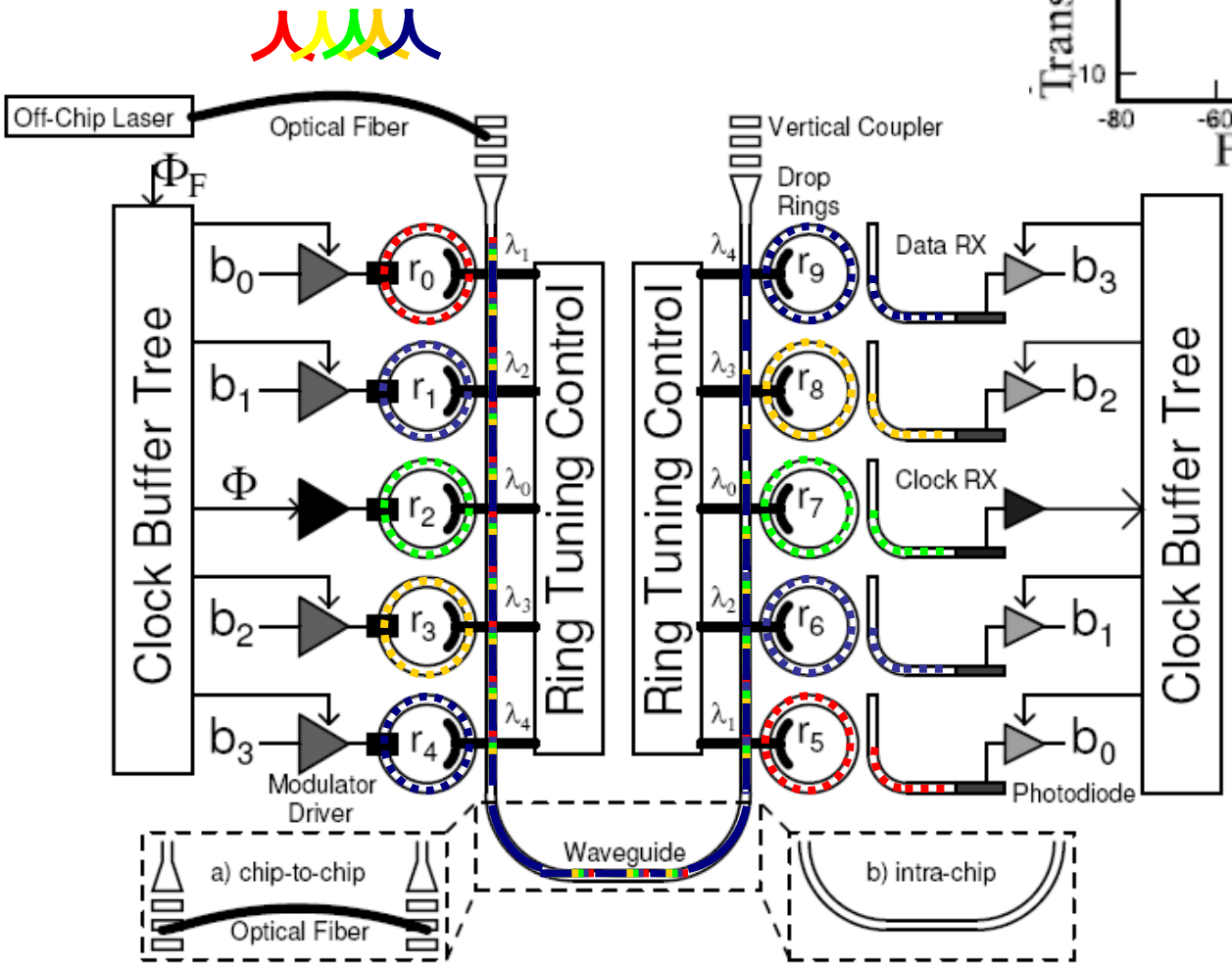
Bulk CMOS Electronics

Si and polySi waveguide formation



Integrated photonic interconnects





- Each λ carries one bit of data

➔Bandwidth Density achieved through DWDM

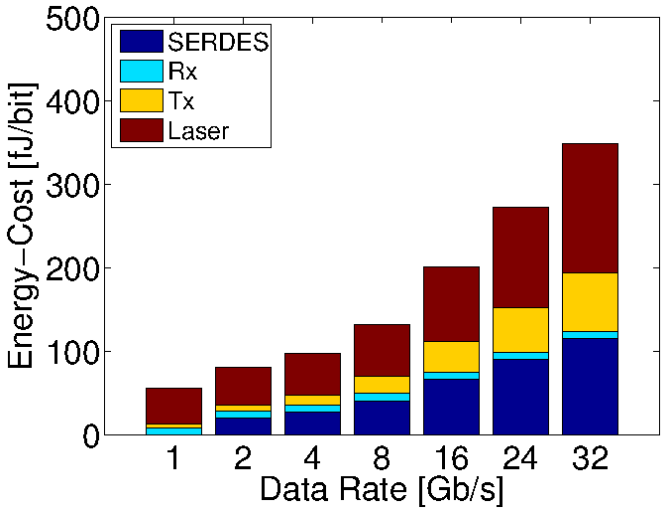
➔Energy-efficiency achieved through low-loss optical components and tight integration

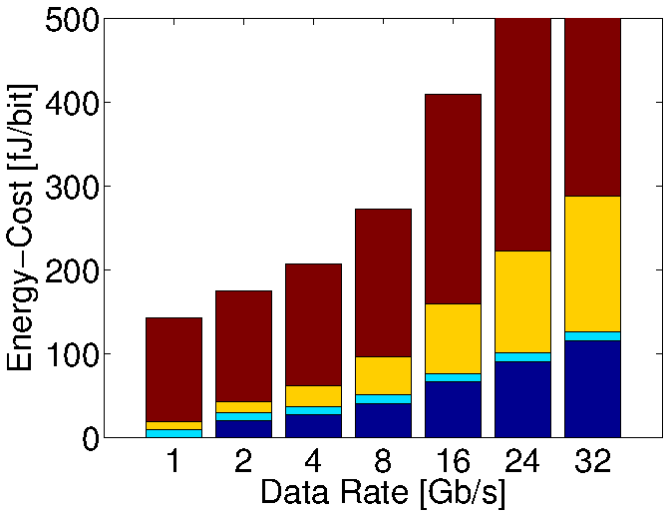
Single channel link tradeoffs

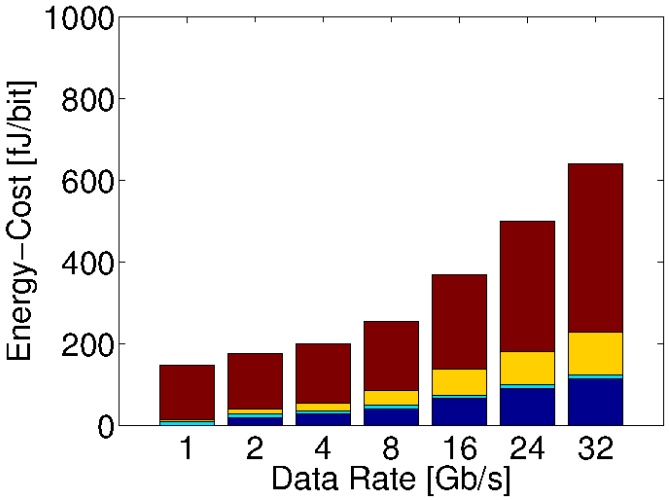
Loss

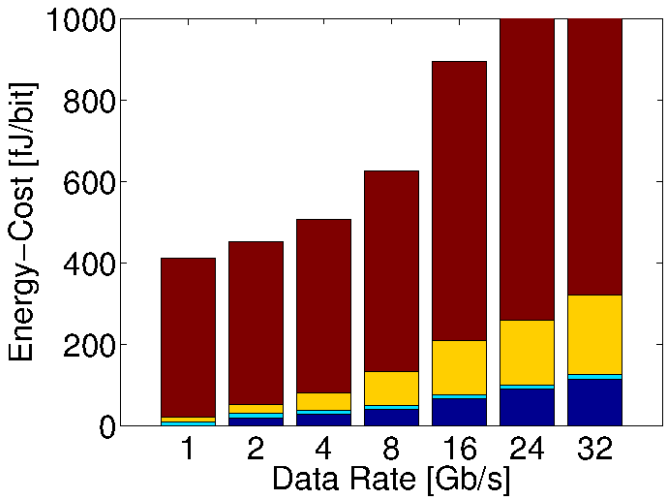
10-dB

15-dB







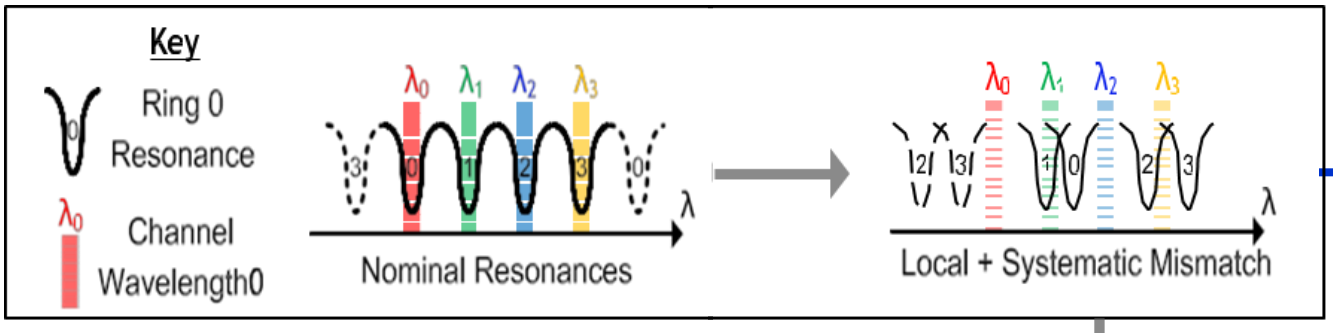


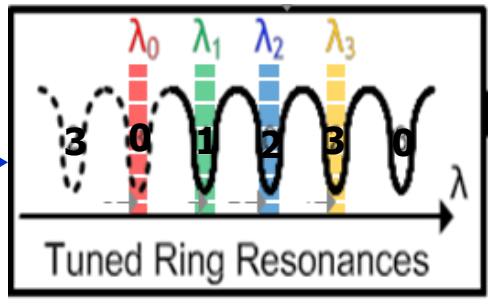
Rx Cap

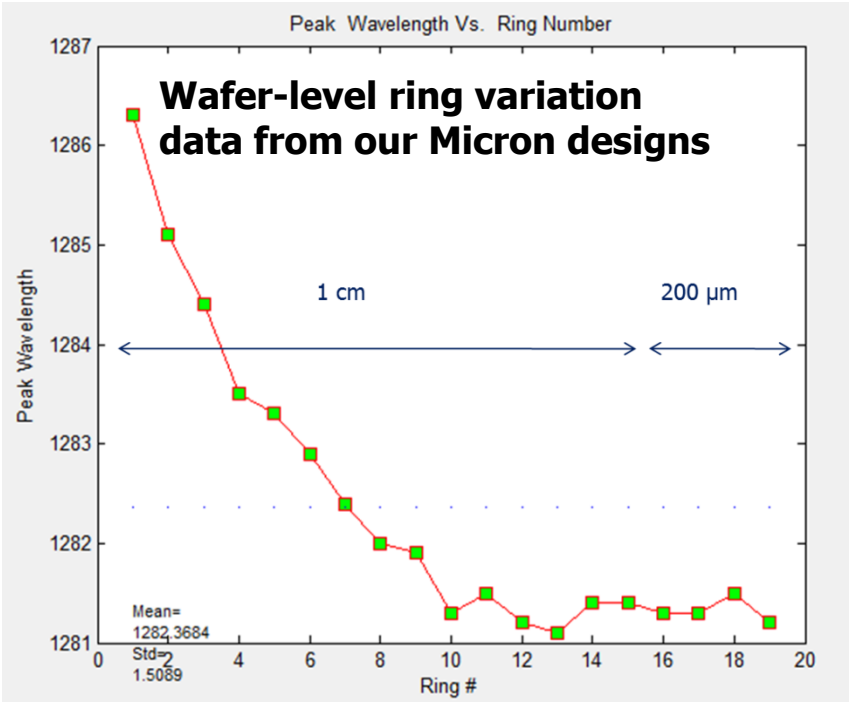
5-fF

25-fF

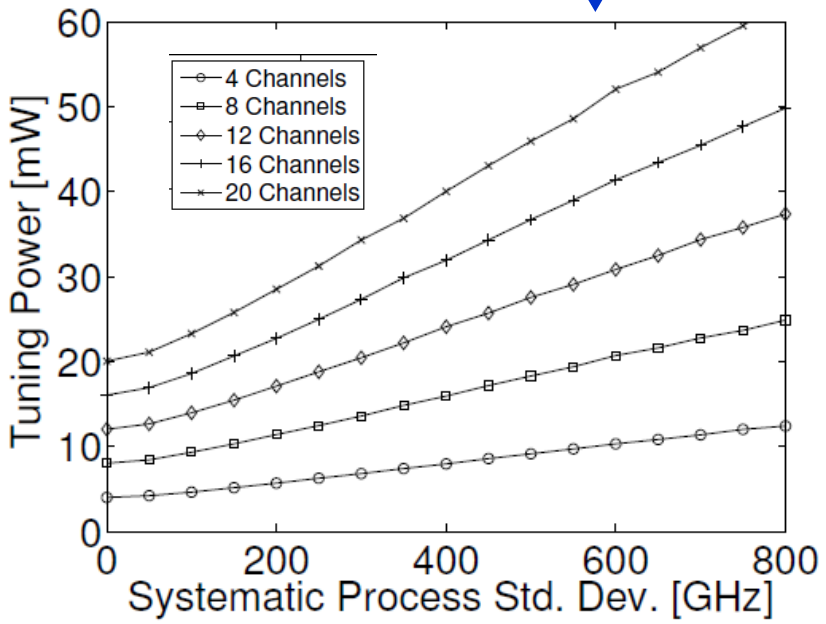
Resonance sensitivity







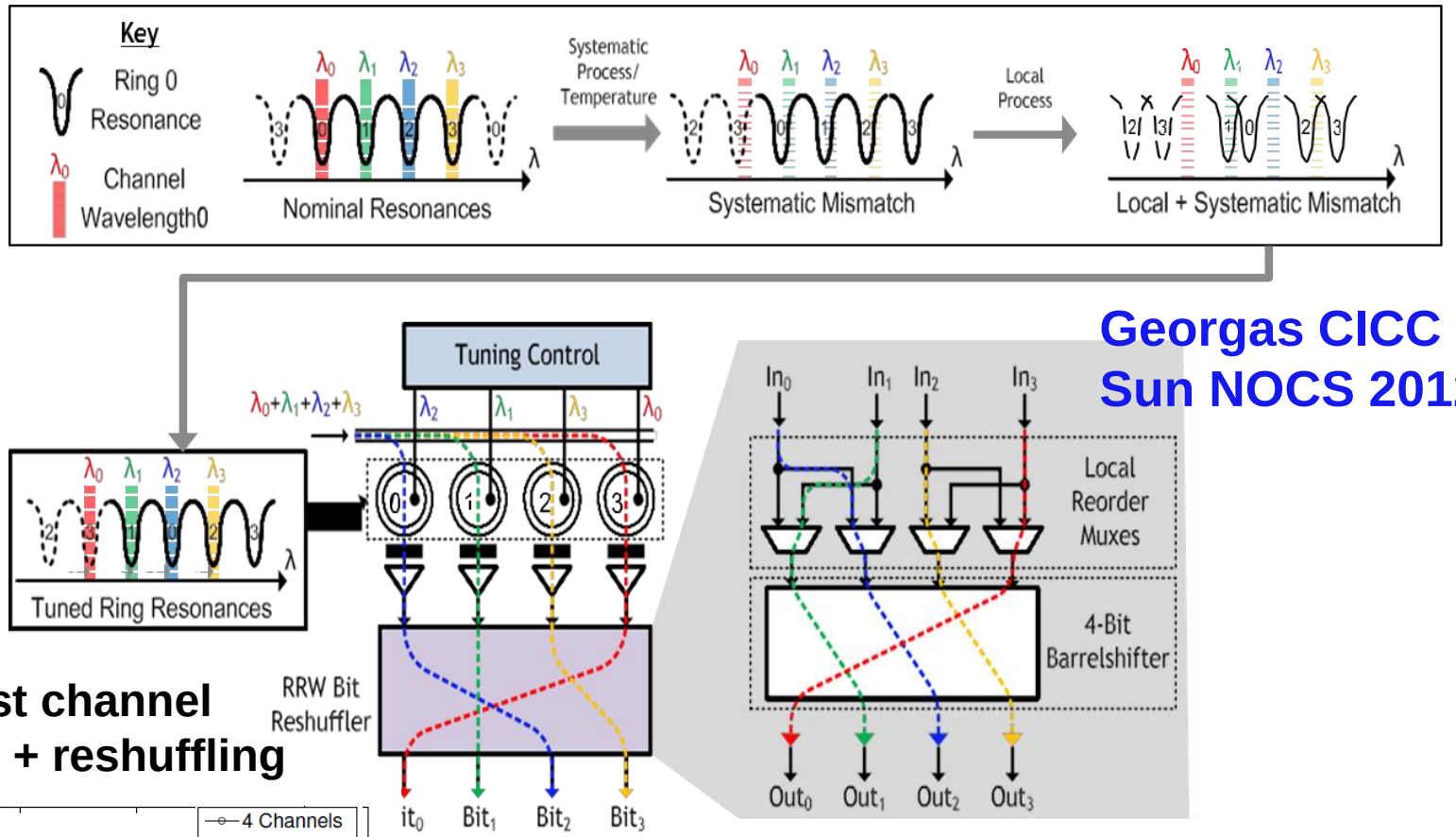
Direct thermal tuning



- Process and temperature shift resonances
- Direct thermal tuning cost prohibitive

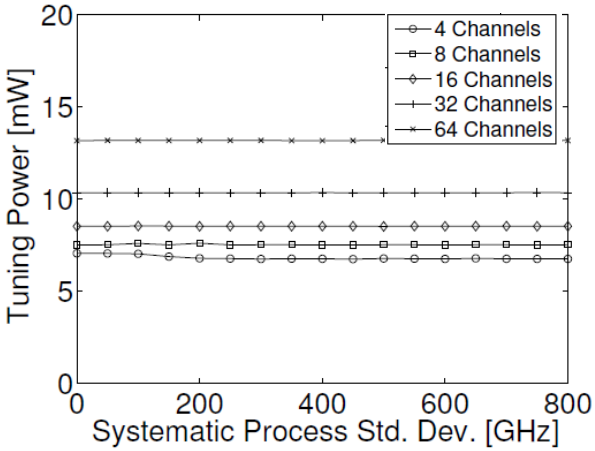
Georgas CICC 2011, Sun NOCS 2012

Smarter wavelength tuning



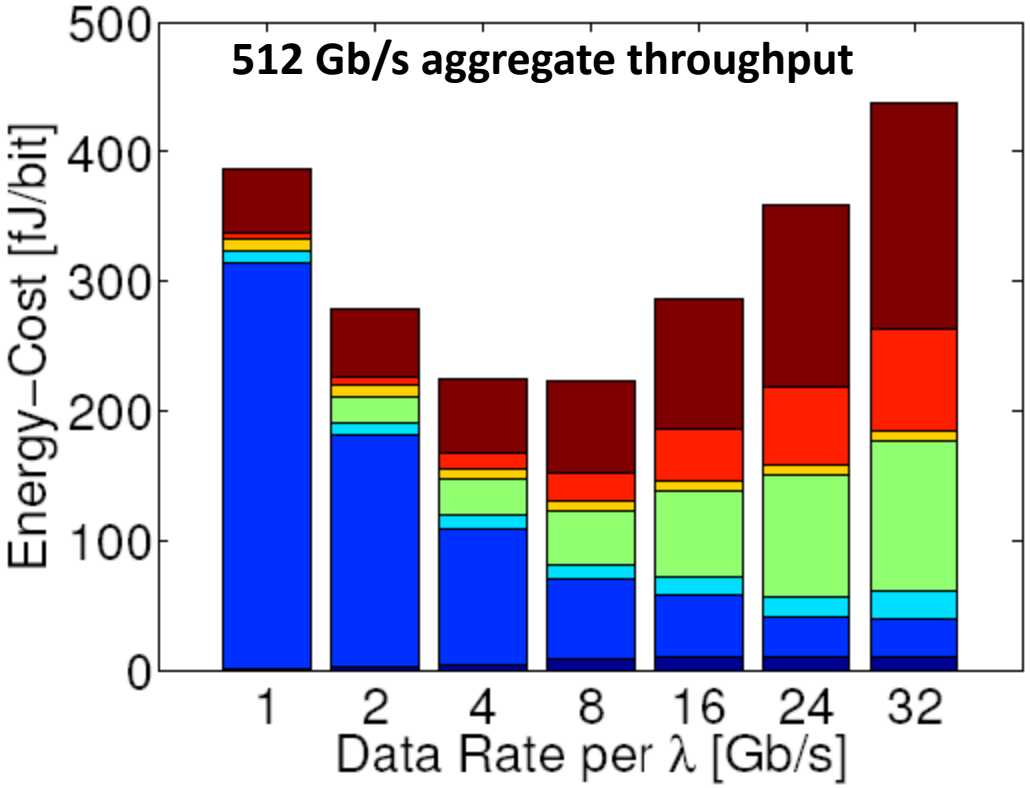
Georgas CICC 2011, Sun NOCS 2012

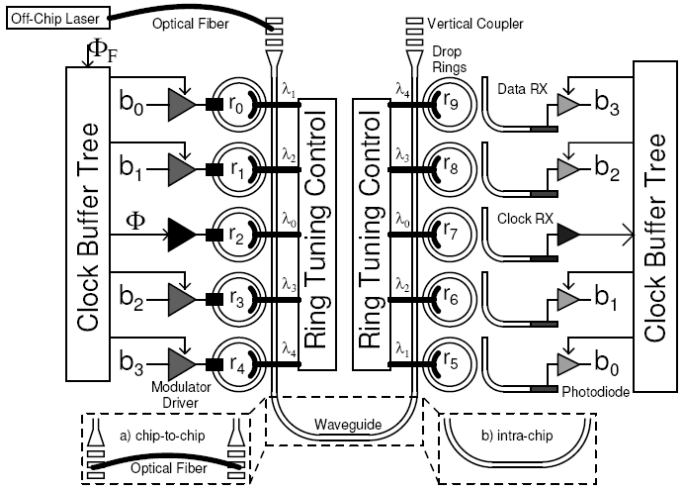
Nearest channel tuning + reshuffling

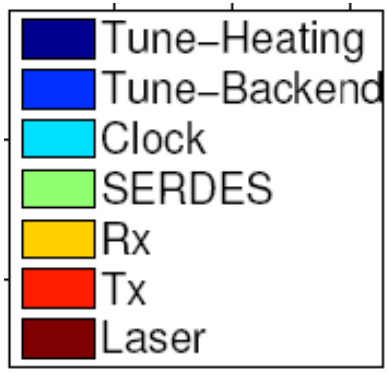


- Utilize systematic global mismatch and temperature shifts
- Electrical backend enables dense WDM
  - Helps reduce tuning costs by more than 10x

Need to optimize carefully





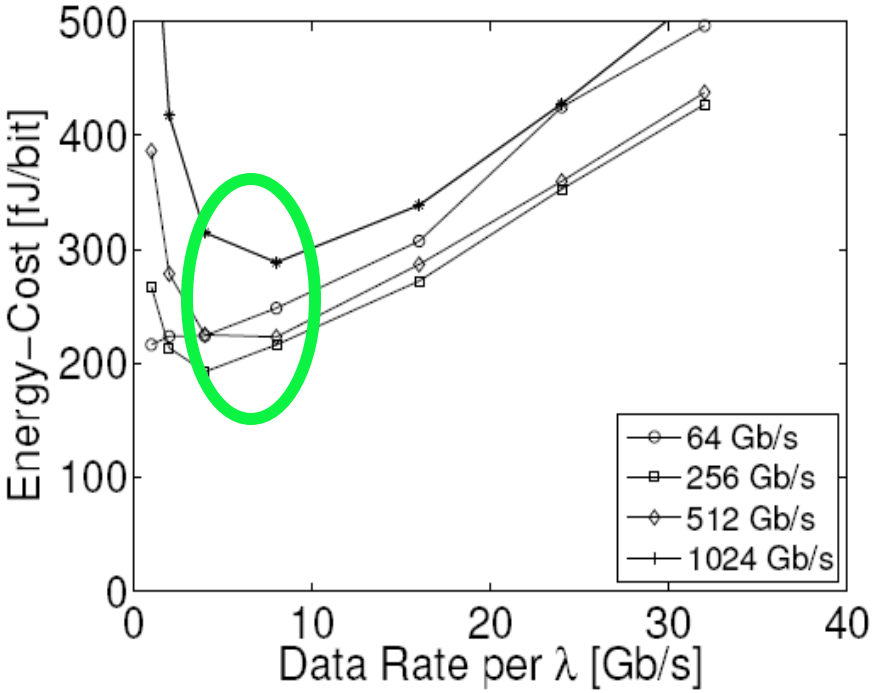


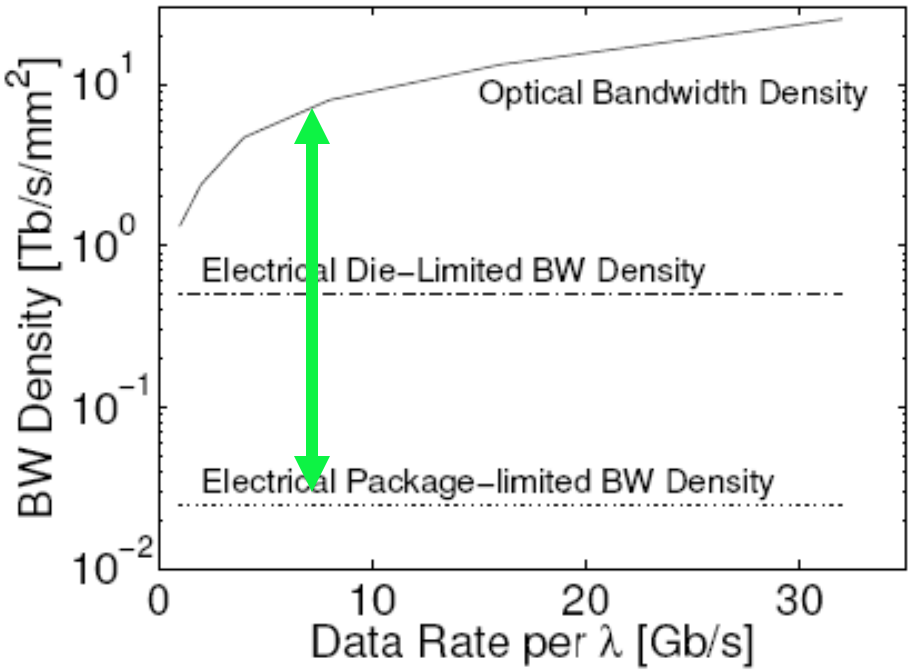
assuming 32nm CMOS

- Laser energy increases with data-rate
  - —Limited Rx sensitivity
  - —Modulation more expensive -> lower extinction ratio
- Tuning costs decrease with data-rate
- **Moderate data rates most energy-efficient**

Georgas CICC 2011

DWDM link efficiency optimization



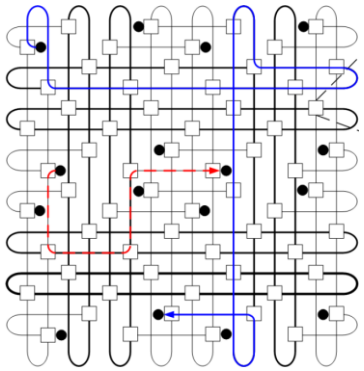


Optimize for min energy-cost

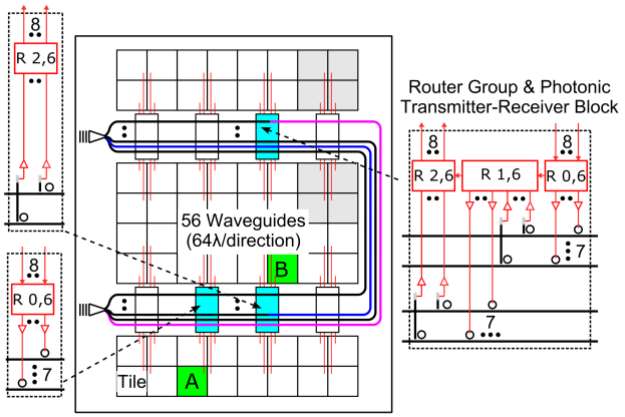
Bandwidth density dominated by circuit and photonics area (not coupler pitch)

- 10x better than electrical bump limited
- 200x better than electrical package pin limit

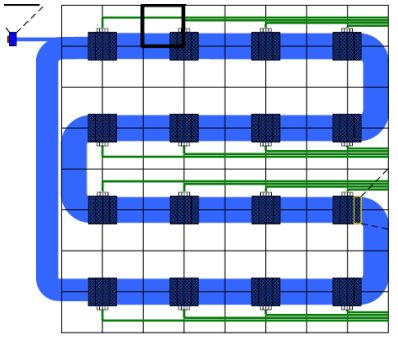
Many architectural studies show promise



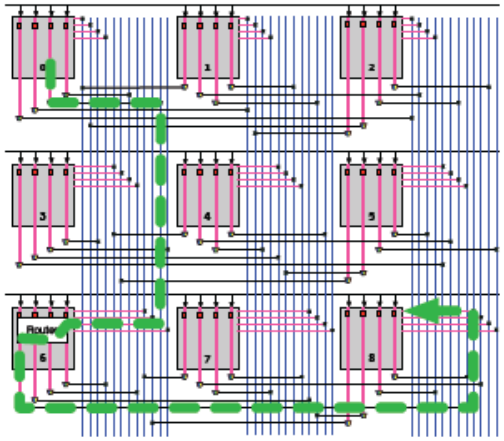
[Shacham'07]
[Petracca'08]



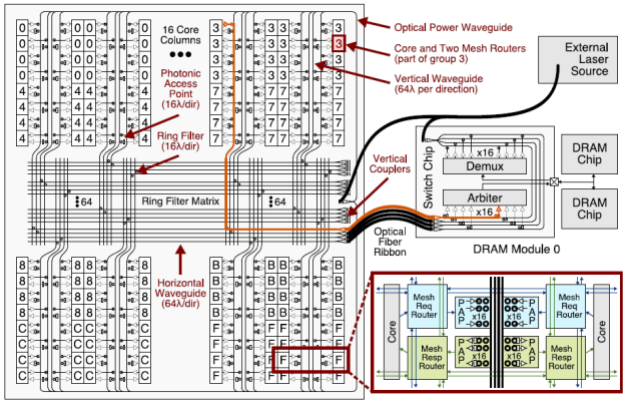
[Joshi'09]
[Pan'09]



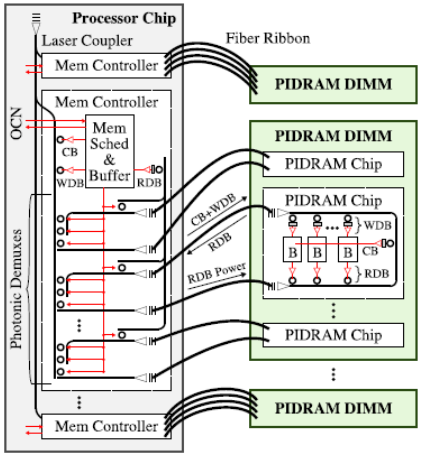
[Vantrease'08]
[Psota'07]
[Kirman'06]



[Koka'08-10]

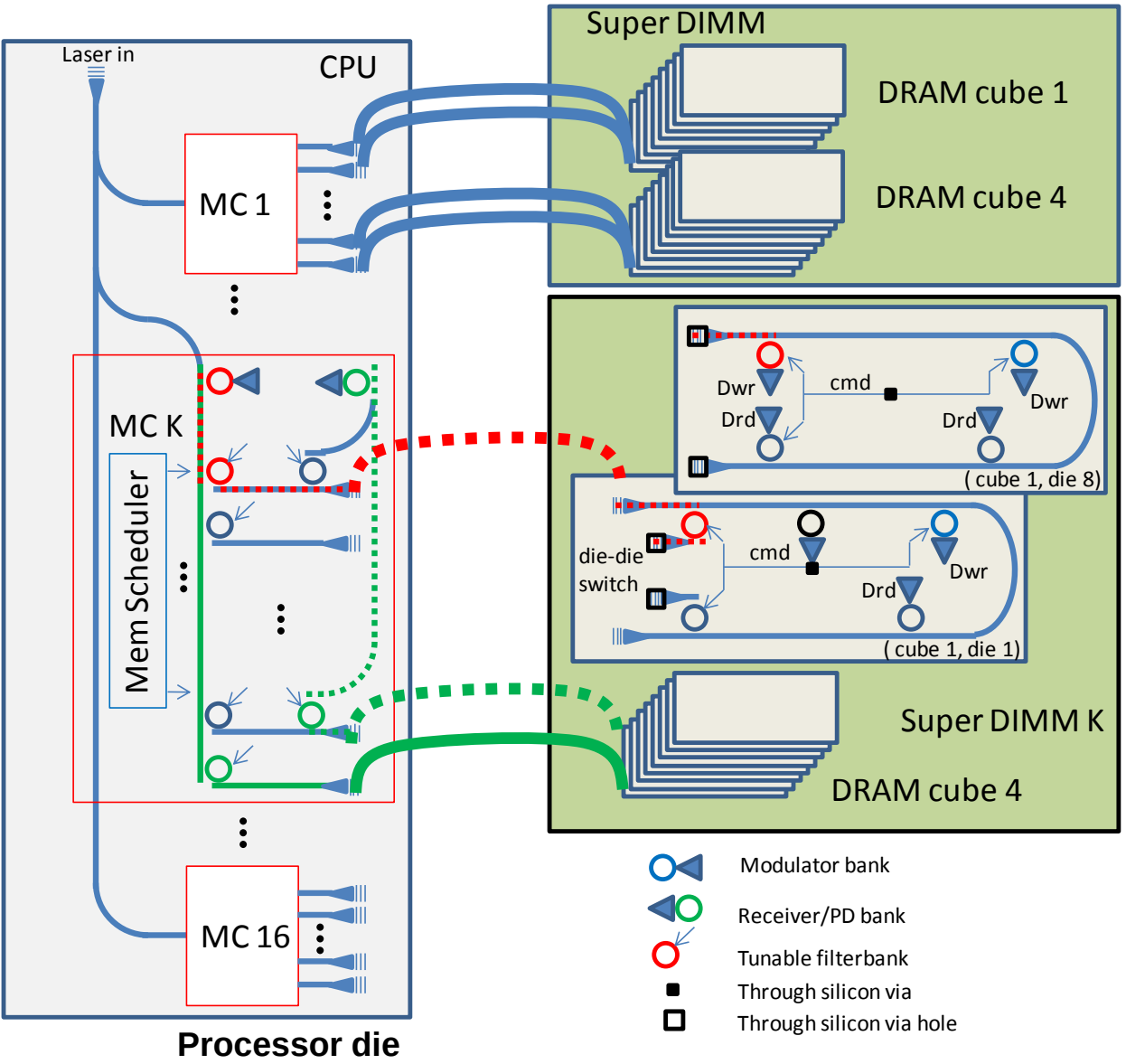


[Batten'08]



[Beamer'10]

Photonic memory interface – leveraging optical bandwidth density



Important Concepts

- Power/message switching (only to active DRAM chip in DRAM cube/super DIMM)

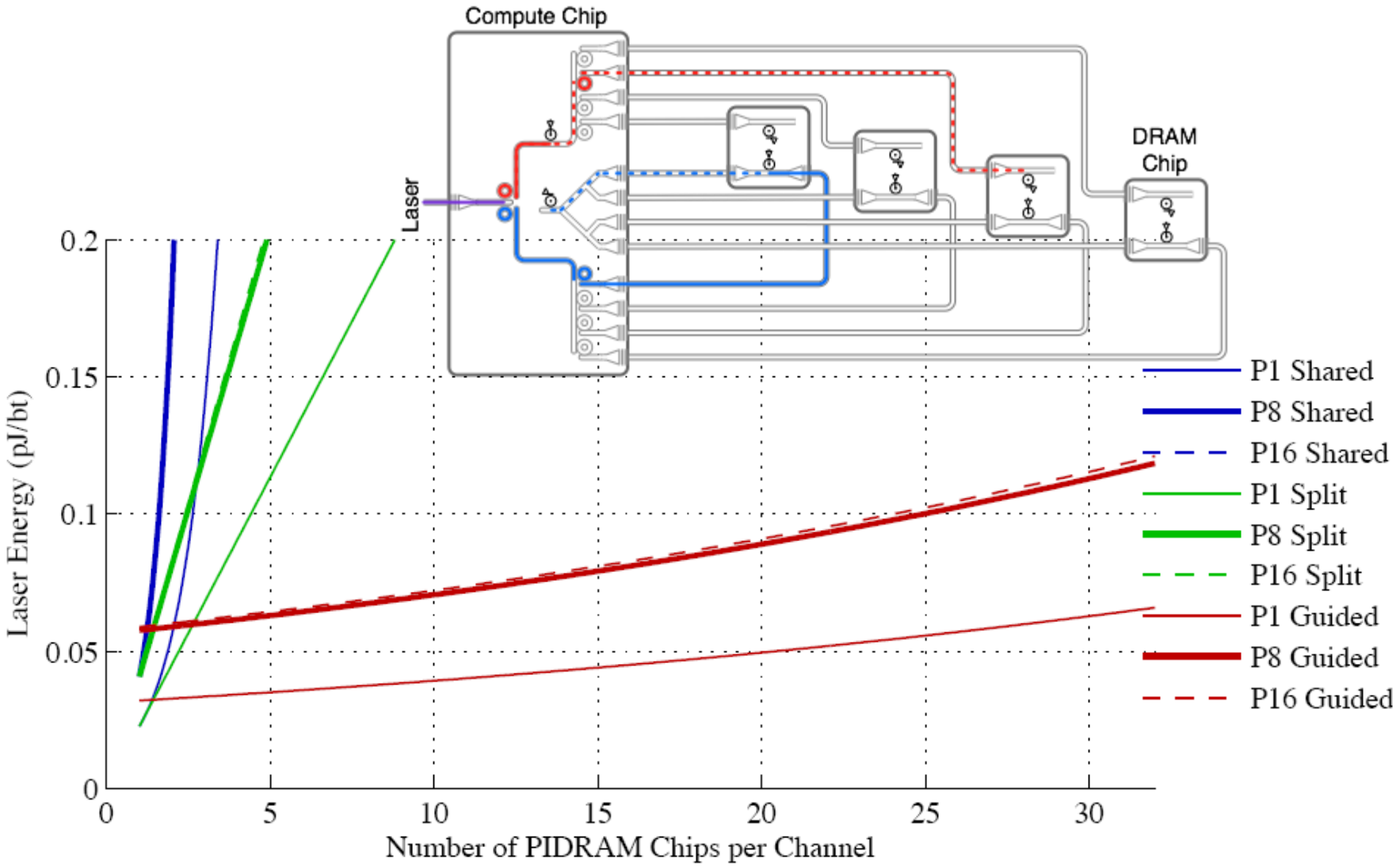
- Vertical die-to-die coupling (minimizes cabling - 8 dies per DRAM cube)

-Command distributed electrically (broadcast)
- Data photonic (single writer multiple readers)

Enables energy-efficient throughput and capacity scaling per memory channel

Beamer ISCA 2010

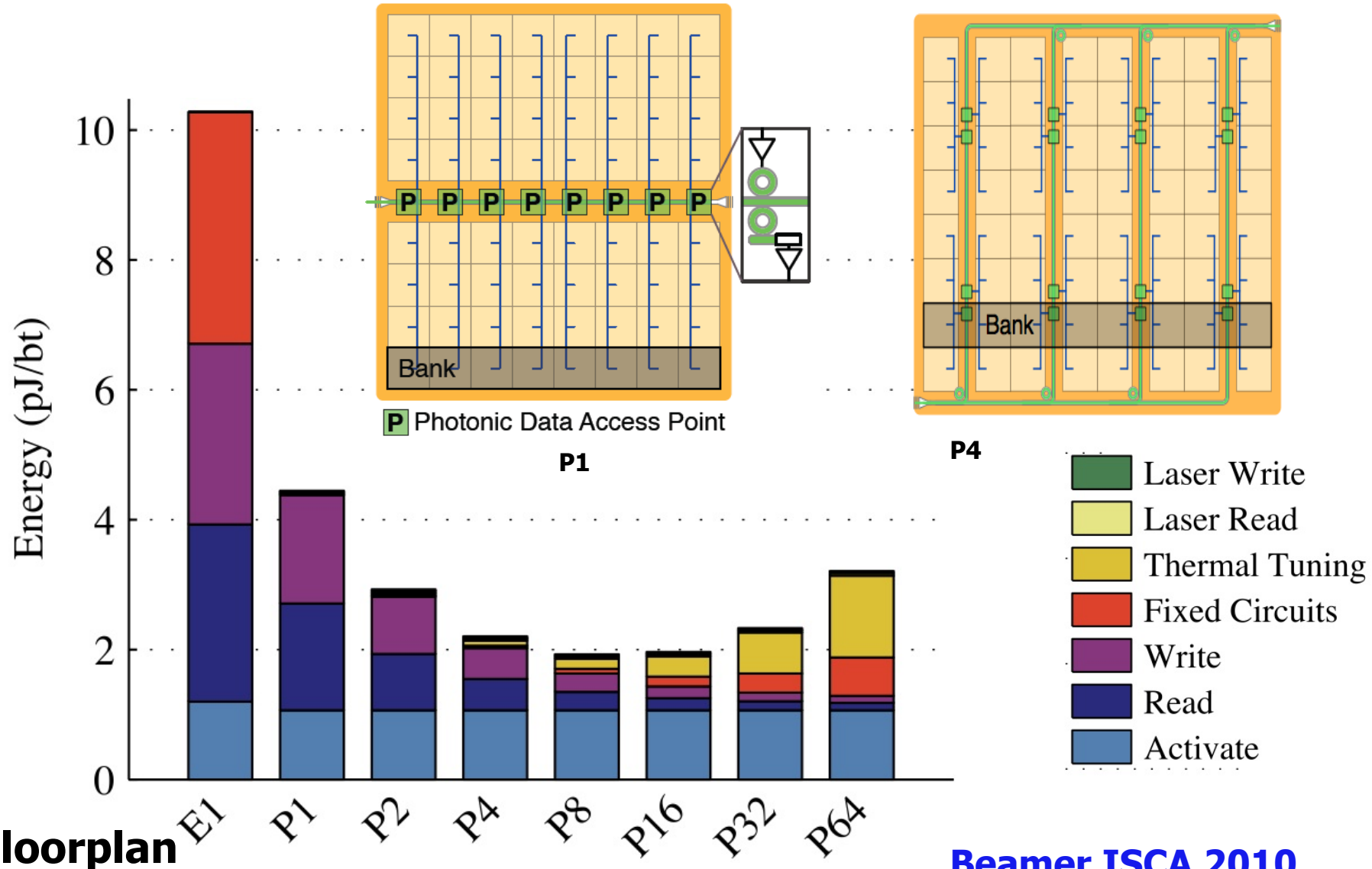
Laser Power Guiding Effectiveness



Enables capacity scaling per channel and significant savings in laser energy

Beamer ISCA 2010

Optimizing DRAM with photonics



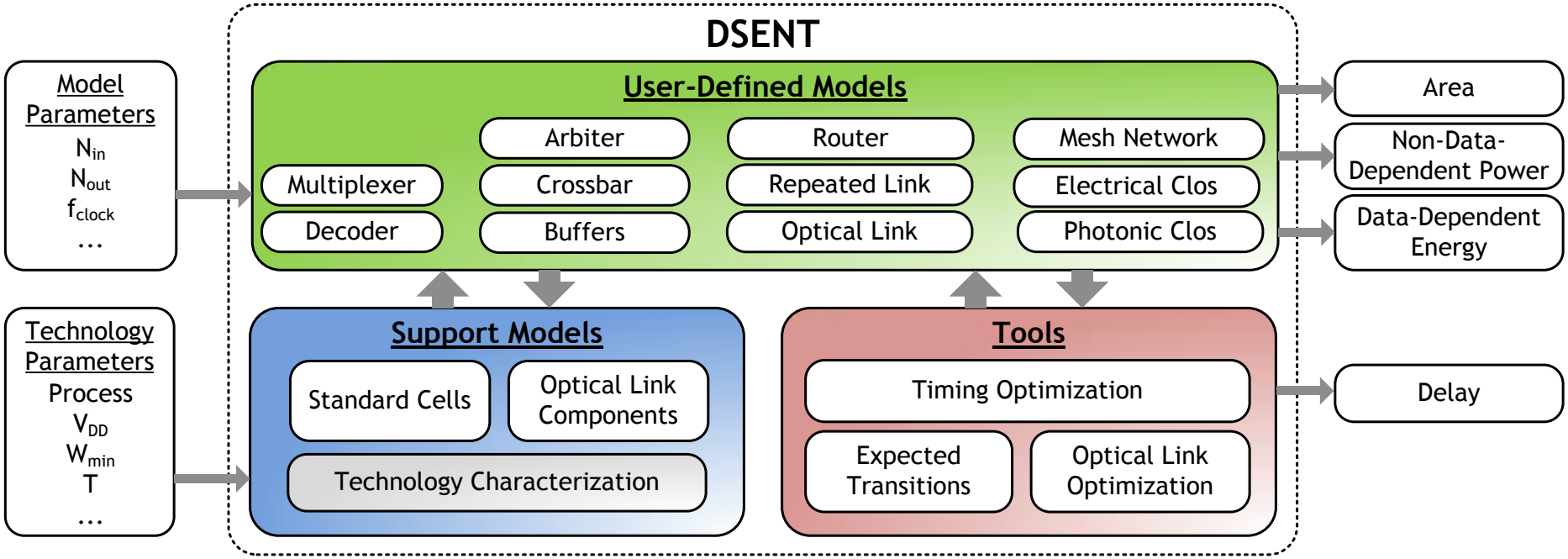
Floorplan

Beamer ISCA 2010

Design Space Exploration of Networks Tool

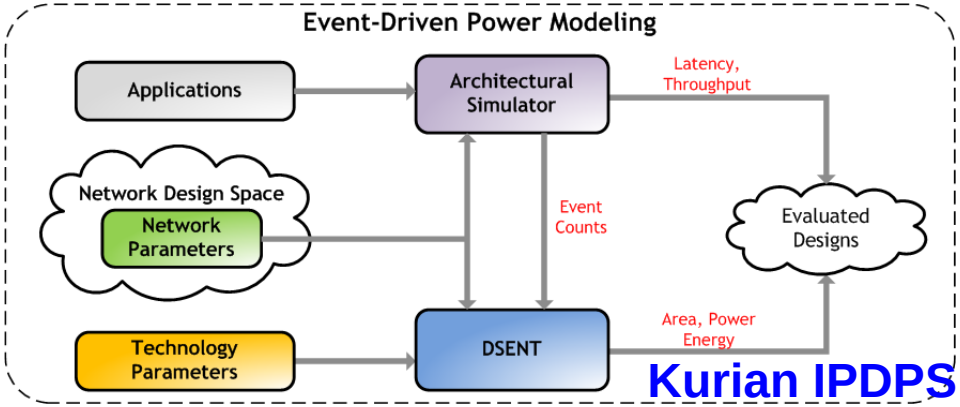
DSENT – A Tool Connecting Emerging Photonics with Electronics for Opto-Electronic Networks Modeling

Chen NOCS 2012



Available for download at:

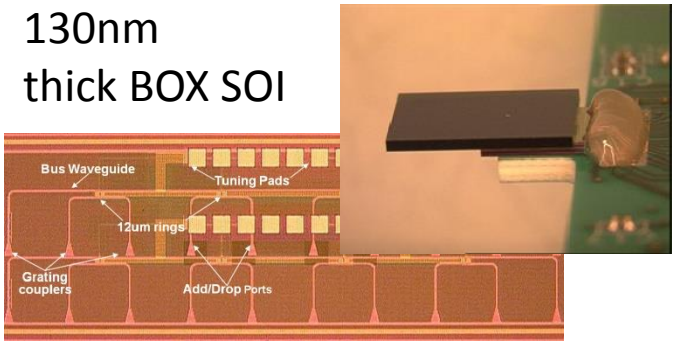
https://sites.google.com/site/mitdsent/



Kurian IPDPS 2012

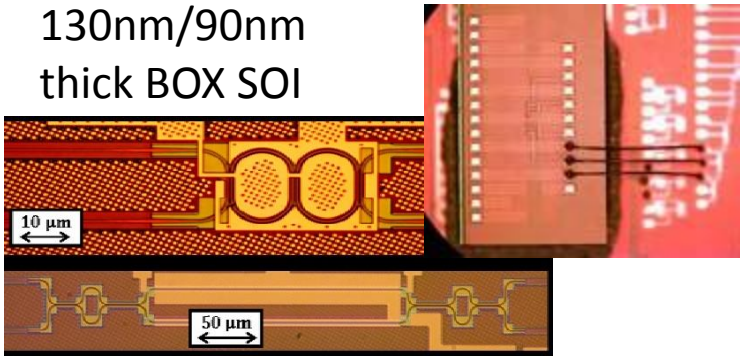
Significant integration activity, but hybrid and older processes …

130nm
thick BOX SOI



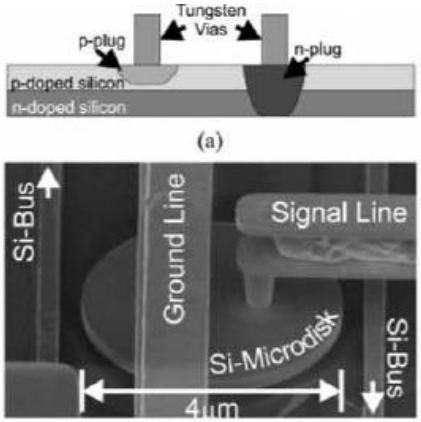
[Luxtera/Oracle/Kotura]

130nm/90nm
thick BOX SOI



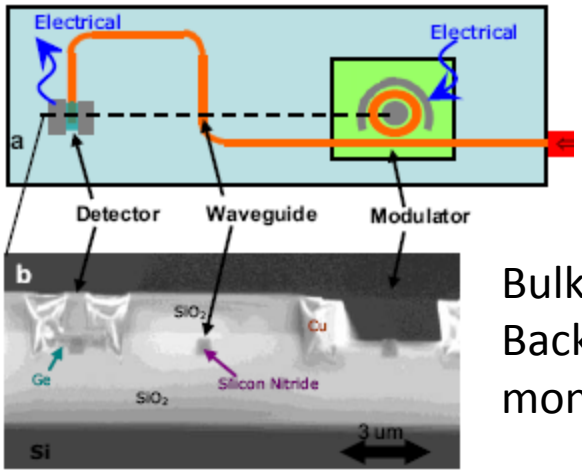
[IBM]

[Many schools]



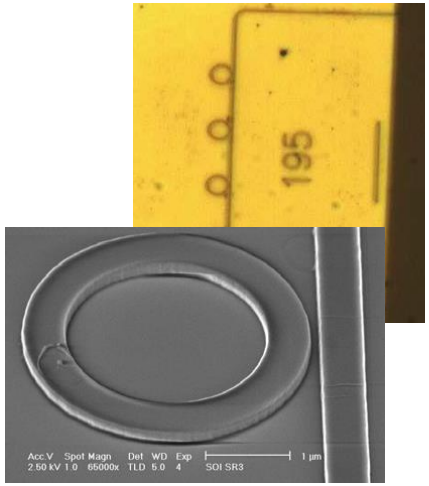
[Watts/Sandia/MIT]

[Lipson/Cornell]



Bulk CMOS
Backend
monolithic

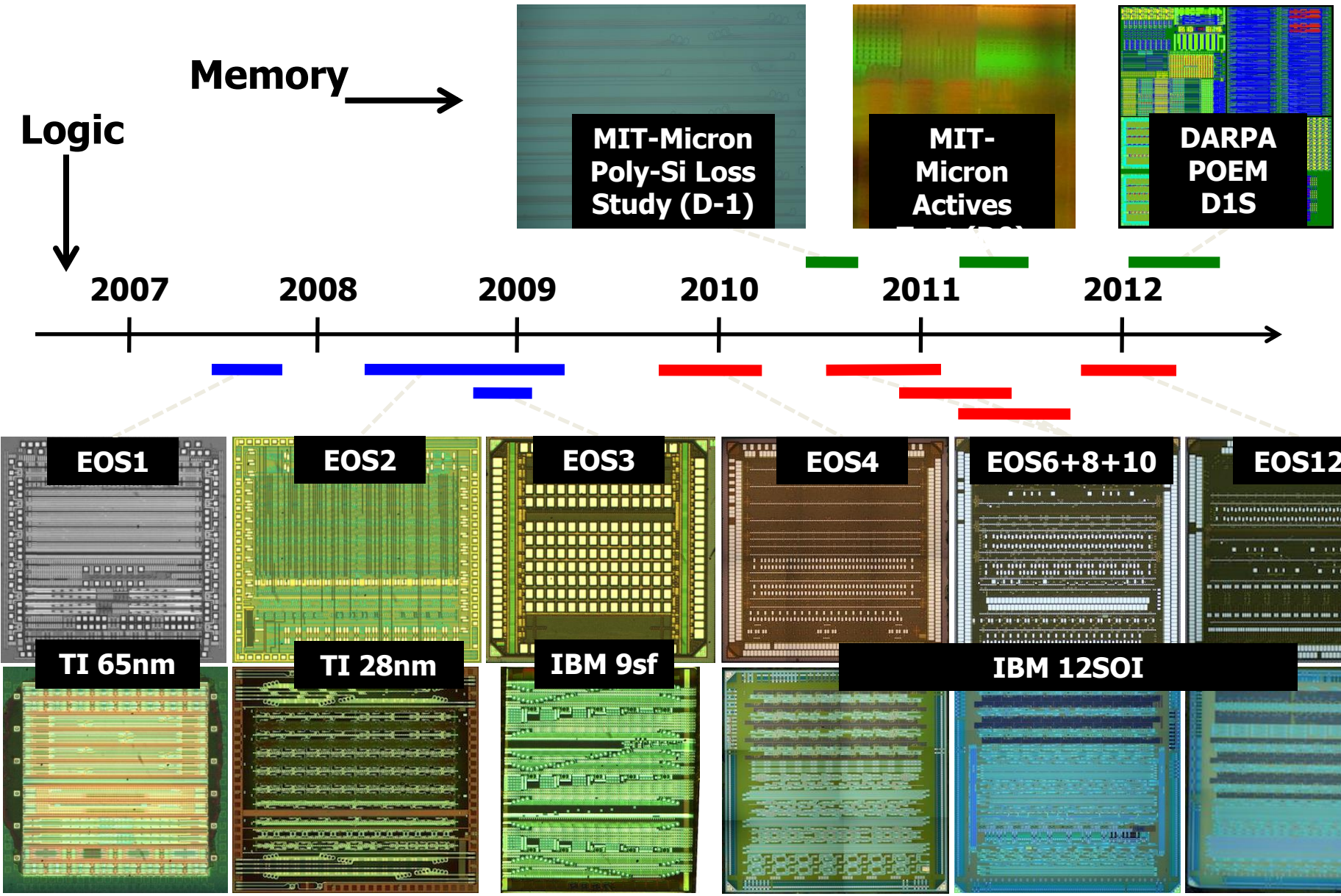
[Intel]



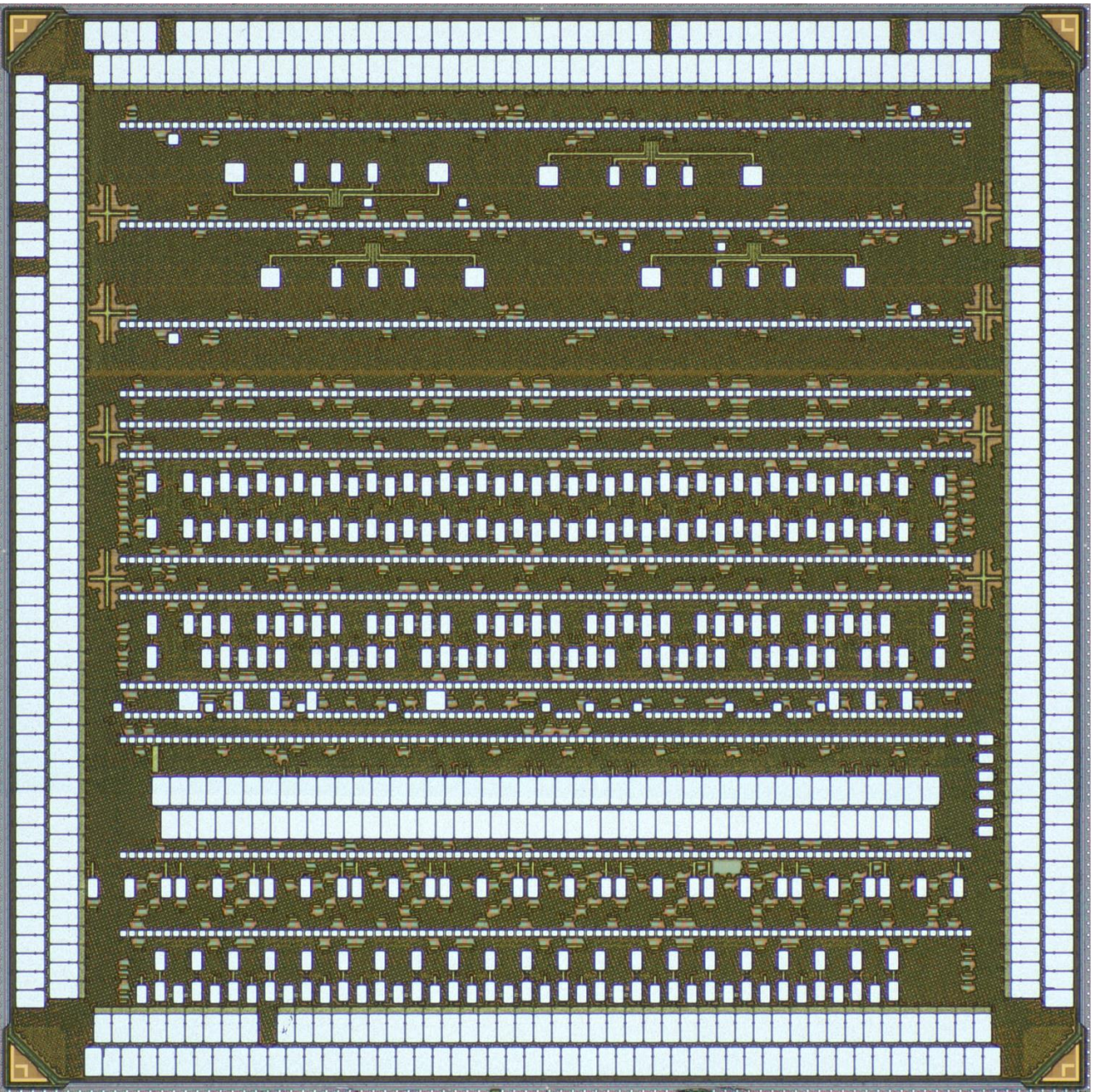
[HP]

[Kimerling/MIT]

Our work: Si Electronic-Photonic Integration Timeline



EOS Platform: EOS8 fabricated in IBM12SOI



Orcutt et al,
Optics Express, 2012

3 x 3 mm die

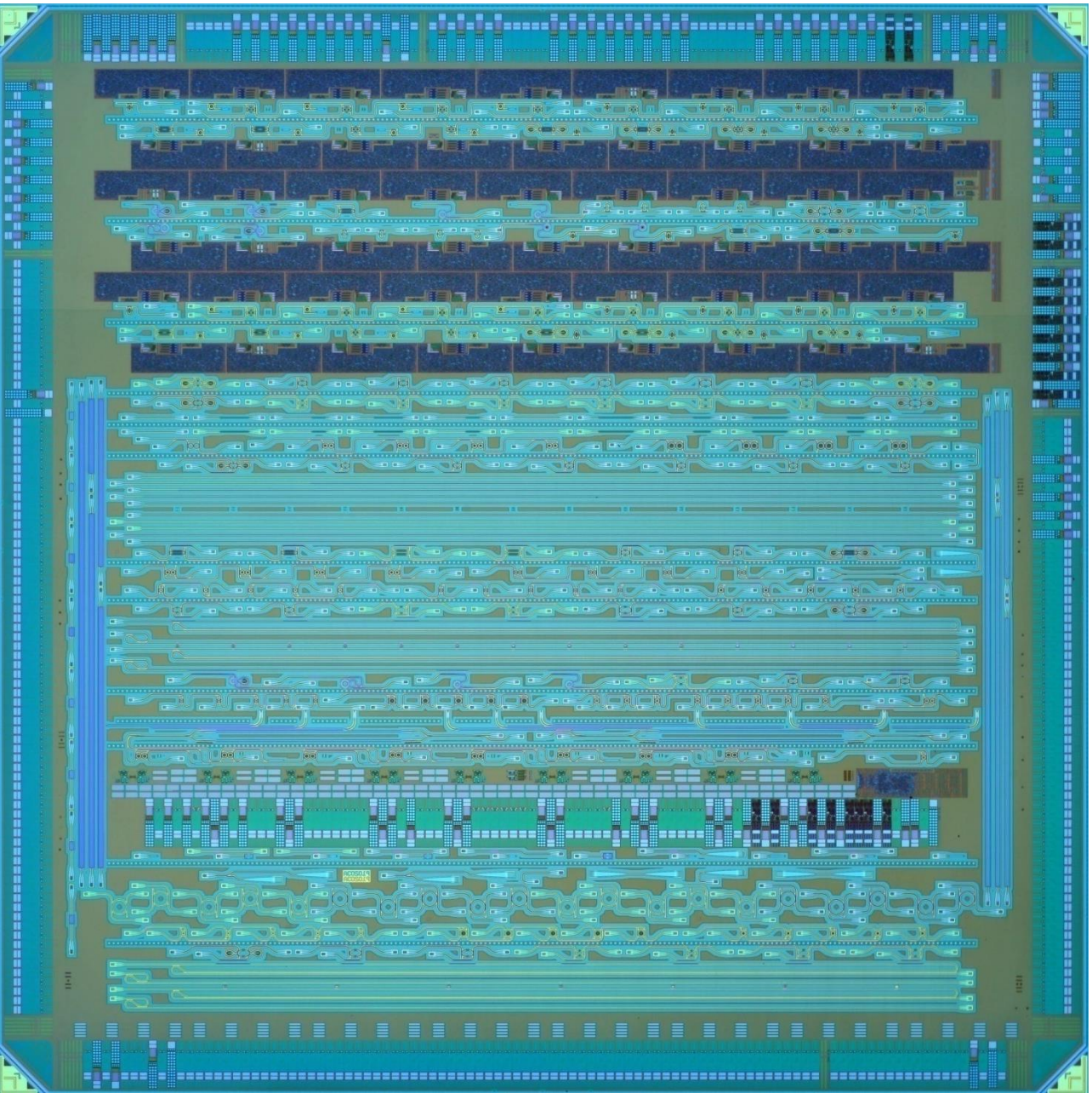
45nm Thin Box SOI Technology
(used for Power 7 and Cell processors)

3M Transistors

400 Pads

ARM Standard Cells and
custom link circuits

EOS8 performance summary



Fiber-to-chip grating couplers with 3.5 dB insertion loss

Waveguides under 4dB/cm propagation loss

10 dB extinction optical modulators

8 channel wavelength division multiplexing filter bank with <-20 dB cross talk

All integrated with electronic circuits

Integration of photonics into VLSI tools

Layout of photonics → abstract → LEF

Layout of Circuit blocks → abstract → LEF

LEF of standard cells, I/O pads

Chip-level verilog (instantiation of .LEF macros and connectivity)

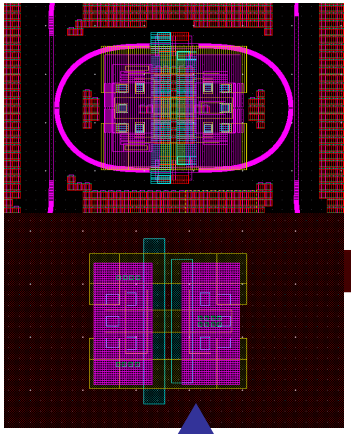
Floorplan (macro placement, power grid, routing Constraints)

Technology files

SOC Encounter Place and route

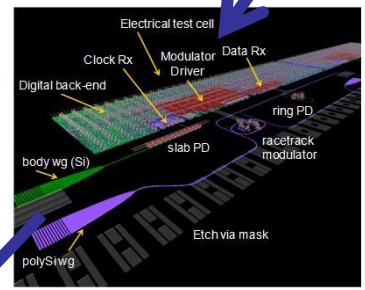
Place&routed layout

layout



modulator.LEF

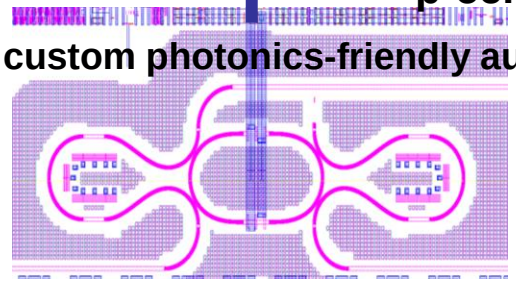
```
VERSION 5.6 ;
BUSBITCHARS "[]" ;
DIVIDERCHAR "/" ;

MACRO block_electronic_etch_row_1
  CLASS BLOCK ;
  ORIGIN -208 -1794 ;
  FOREIGN block_electronic_etch_row_1  208 1794 ;
  SIZE 2488 BY 165 ;
  SYMMETRY  X Y R90 ;
  PIN heater_a_1
    DIRECTION INOUT ;
    USE SIGNAL ;
    PORT
      LAYER ua ;
        RECT 431 1870.5 436.5 1882 ;
    END
  END heater_a_1
...
  OBS
    LAYER m1 ;
      RECT 208 1794 2696 1959 ;
...
  END
END block_electronic_etch_row_1

END LIBRARY
```

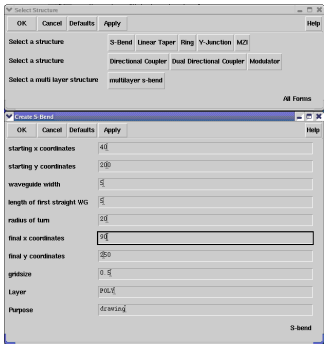


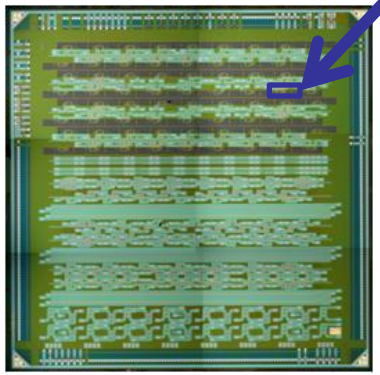
abstract

Photonic device p-cell

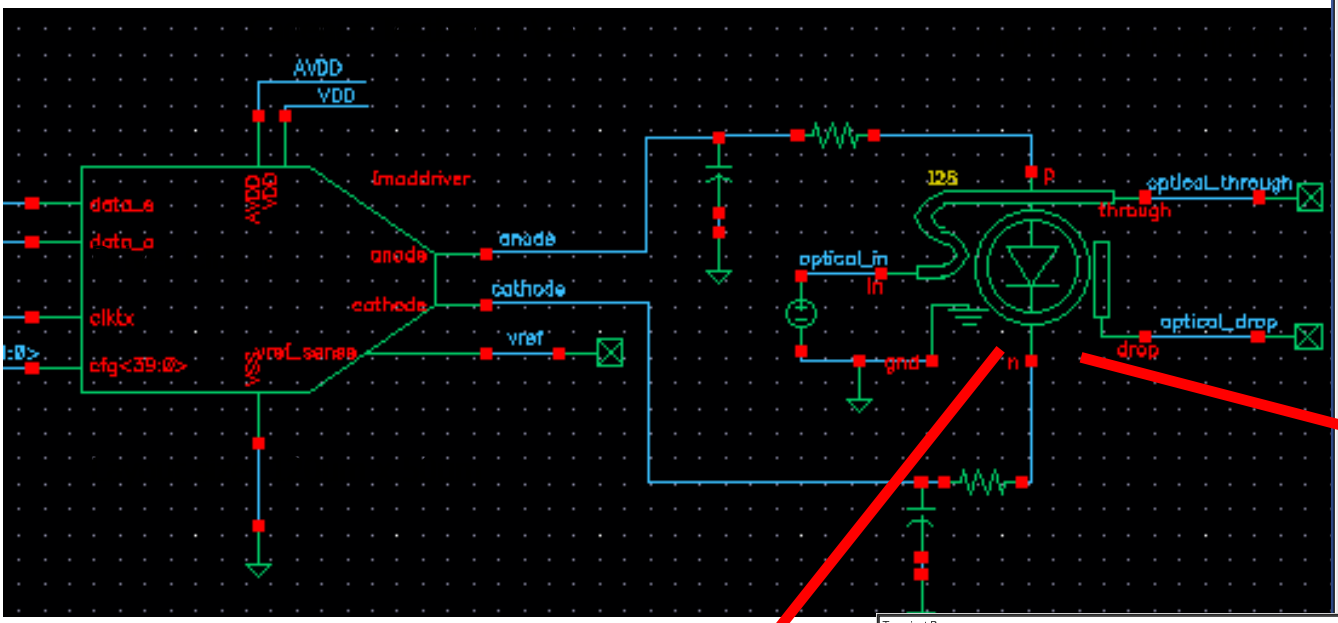
custom photonics-friendly auto-fill





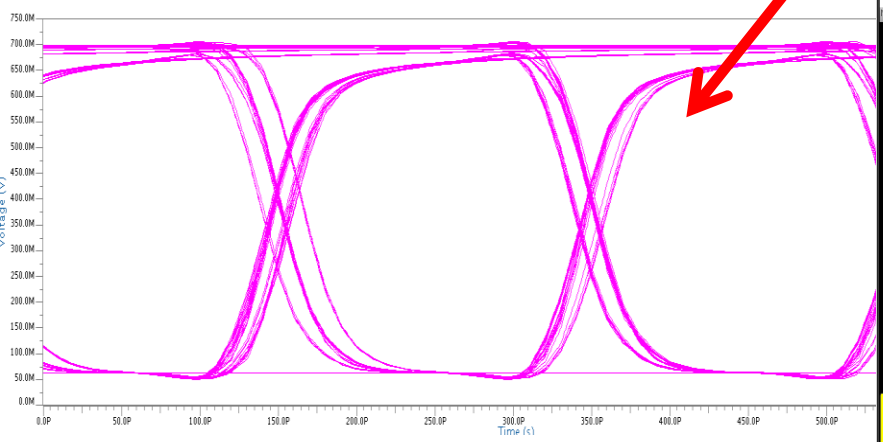


Circuit/Device Co-Simulation: VerilogA

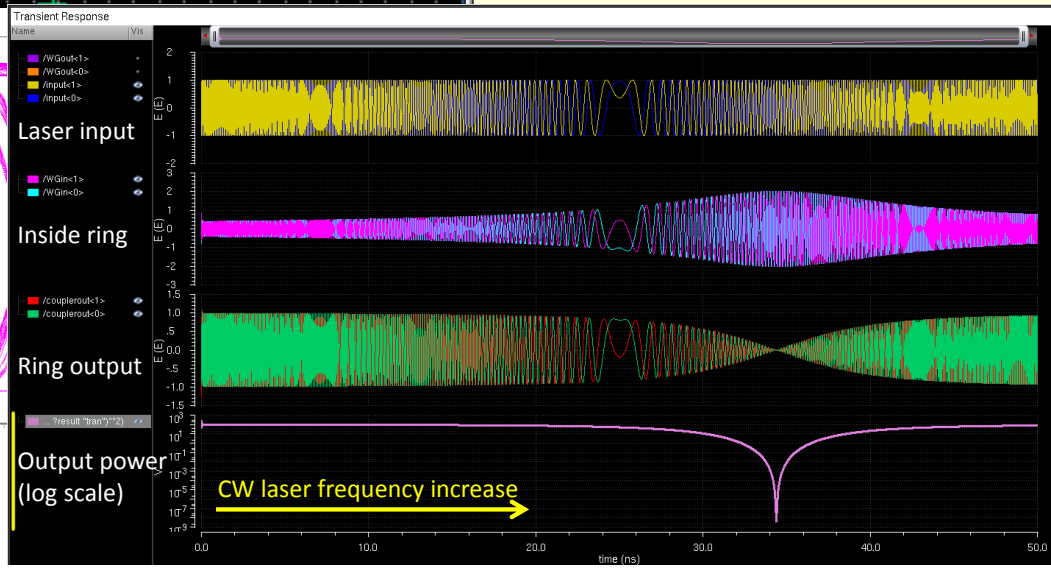


Verilog A Model

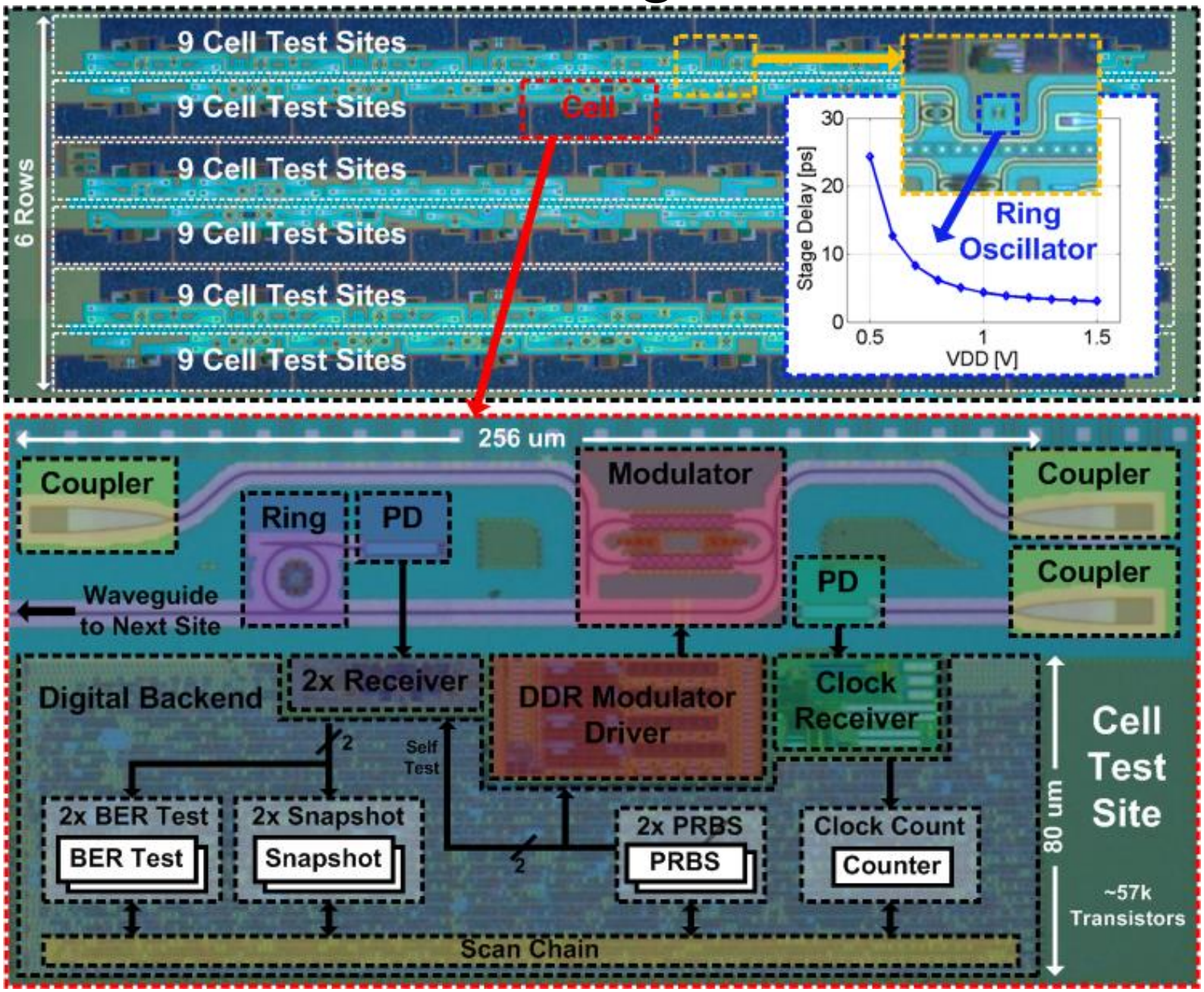
```
compute-0-12: /home/isg_chips/projects/micron_d1/chip/d1_moddriver/modulat -
Session  Edit  View  Bookmarks  Settings  Help

// VerilogA for d1_moddriver, modulator_ridge, veriloga

`include "disciplines.vams"

/*
Changelog:

22 Dec Updates from Milos & Jeff:
*/
module modulator_ridge(p, n, gnd,

        //-------------------------
        // DEFINE PORTS
        //-------------------------

        inout p, n, gnd;
        input in;
        output through, drop;
        electrical p, n, gnd, in, through, drop;
        branch (p,n) res, cap;


        //-------------------------
        // OPTICAL PARAMETERS
        //-------------------------

        // Physical Constants

        parameter real pi       = 3.1415926536;         // Pi [unitless]
        parameter real k        = 1.38e-23;             // Boltzmann const
        parameter real T        = 300;                  // Temperature [K]
        parameter real q        = 1.6e-19;              // Charge of elect
        parameter real eo       = 8.85e-12;             // Permittivity of
        parameter real es       = 12;                   // Relative permit
        parameter real c        = 3e8;                  // Speed of light
        parameter real nf       = 3e-27;                // Refractive inde
```



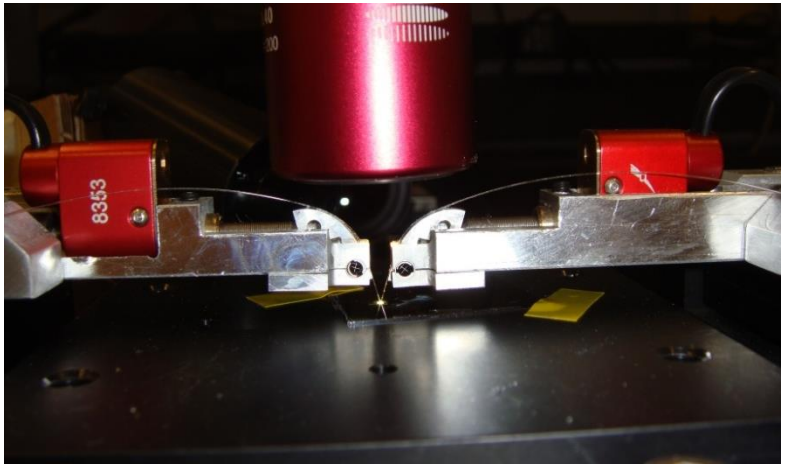
Optical Eye Diagram

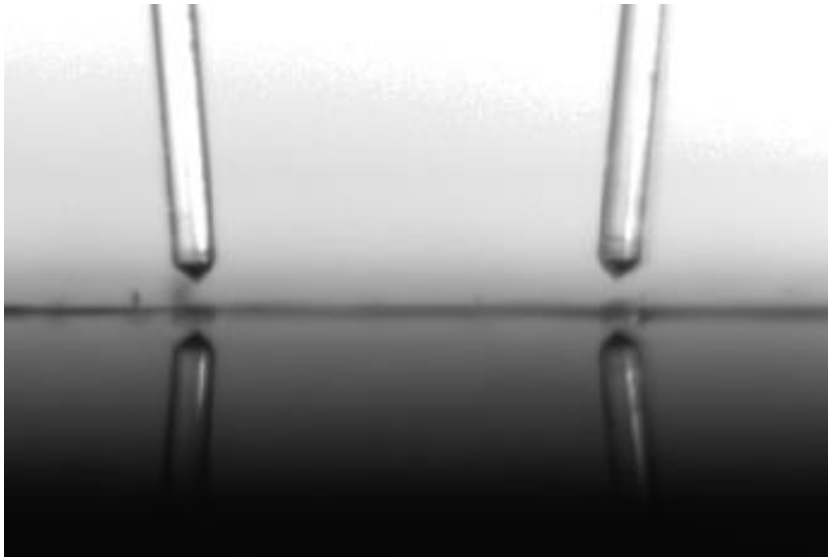


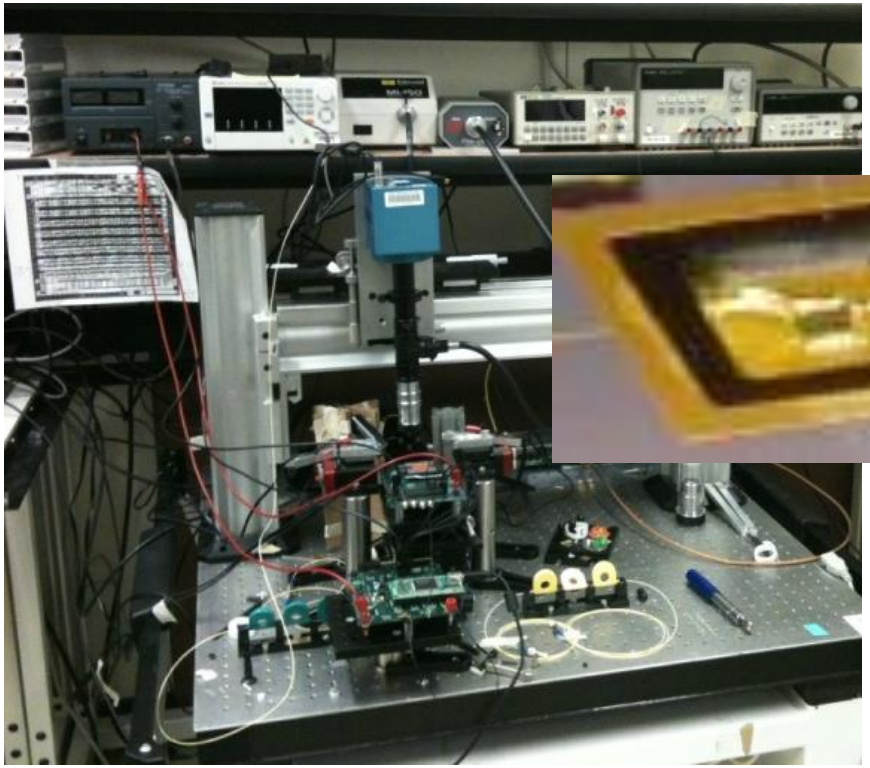
Platform Organization

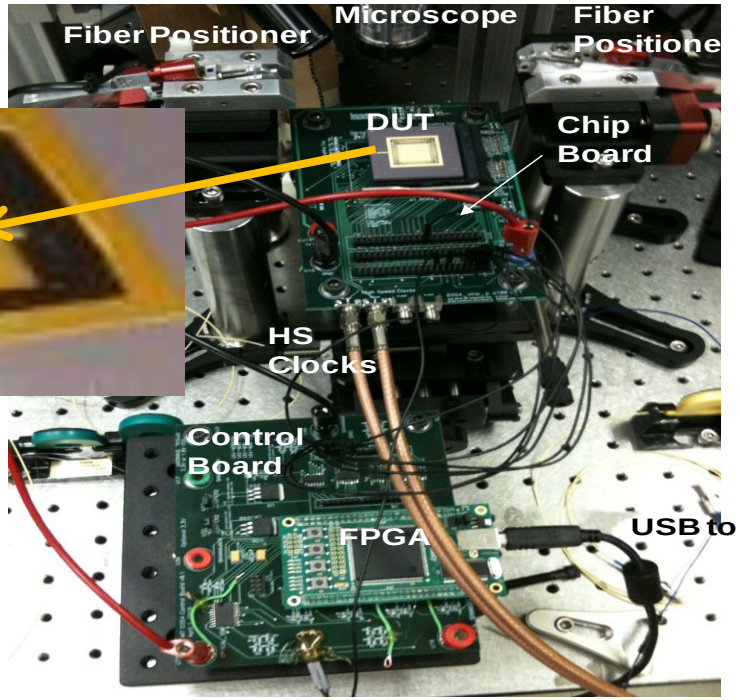


Chips fully packaged



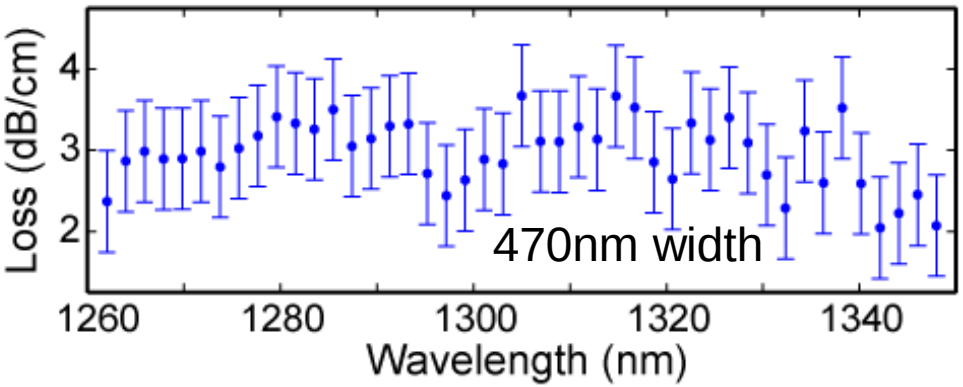


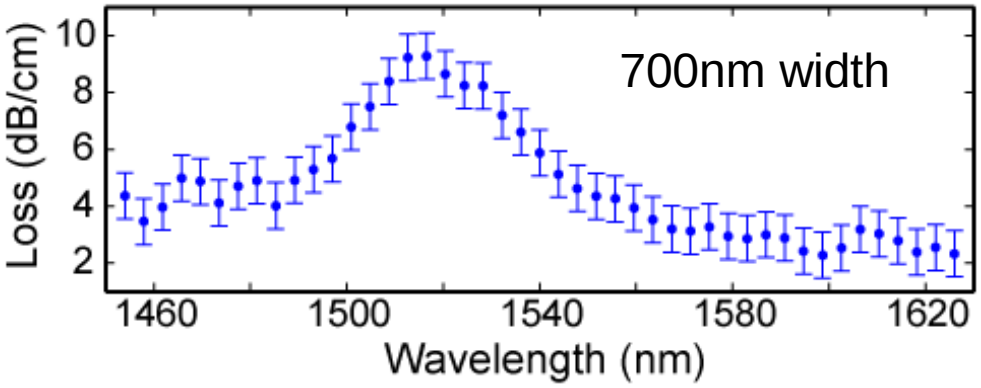




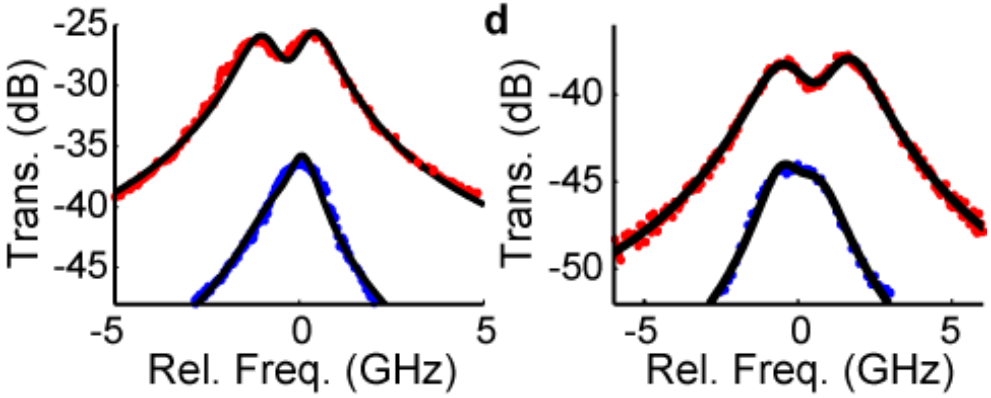
Best waveguide losses ever reported in a sub-100nm production CMOS line

- Body-Si waveguides
  - 3-4dB/cm loss
- Poly waveguides
  - 50dB/cm loss

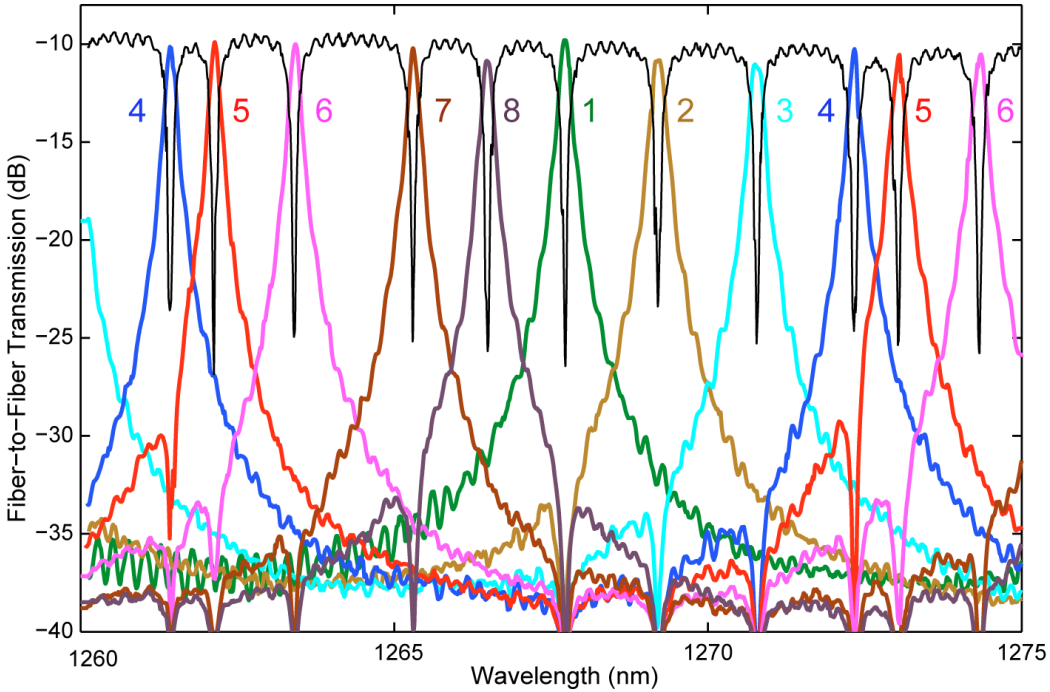


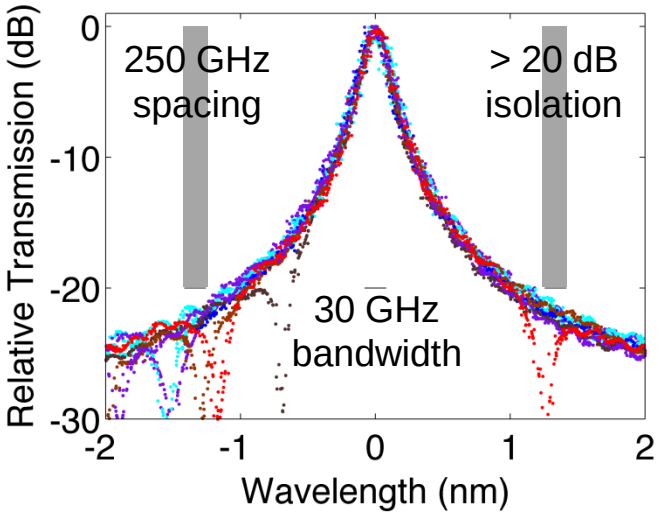


- Body-Si ring Q factor
  - 227k @ 1280nm
  - 112k @ 1550nm



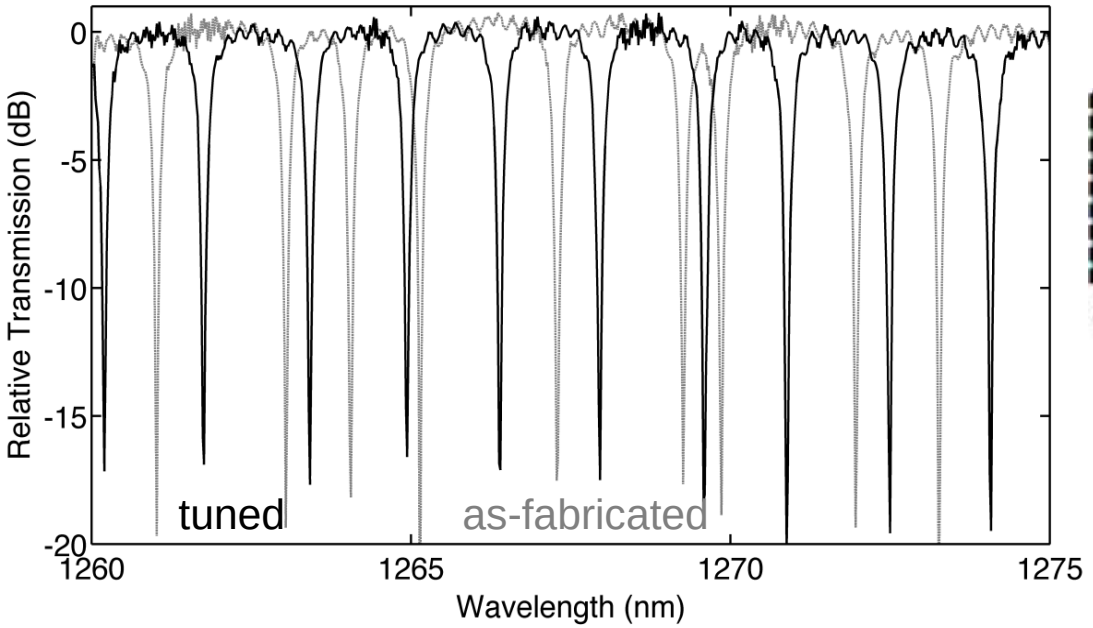
Exceptional dimensional control in 45nm node



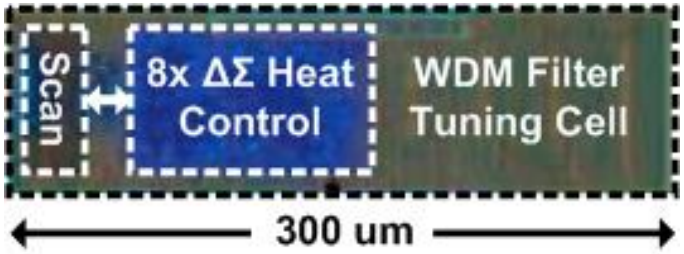


- 8-wavelength filterbank results
  - Filter channels fabricated in order
  - Less than 1nm variation
- Excellent channel isolation (>20dB at 250GHz spacing)

Integrated thermal tuning circuits

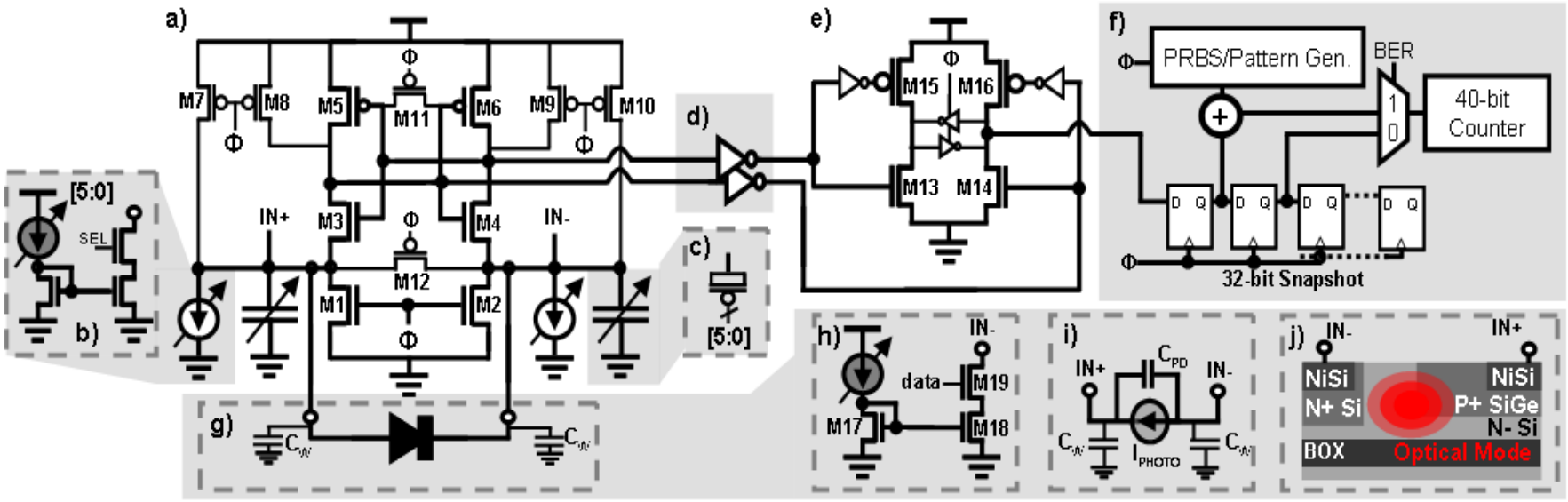


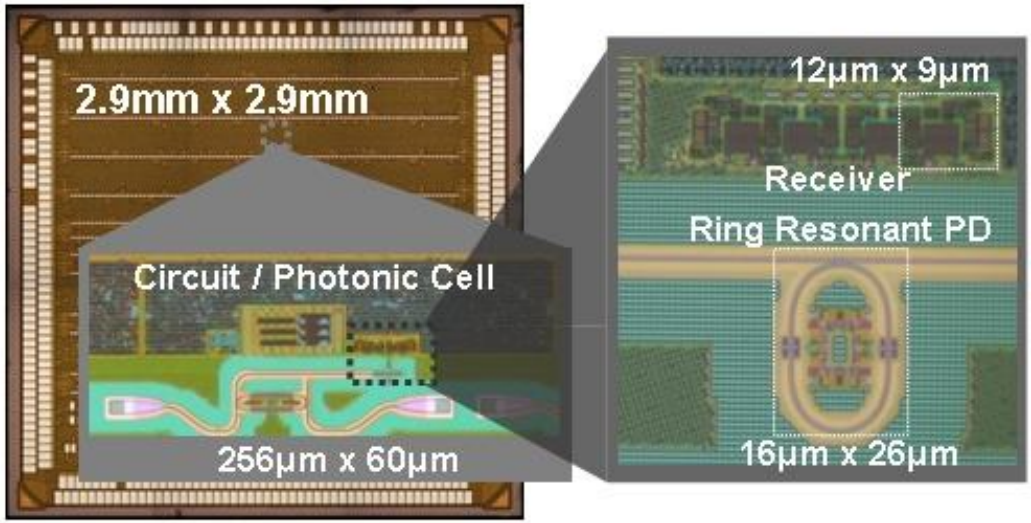
integrated digital PWM heater controller

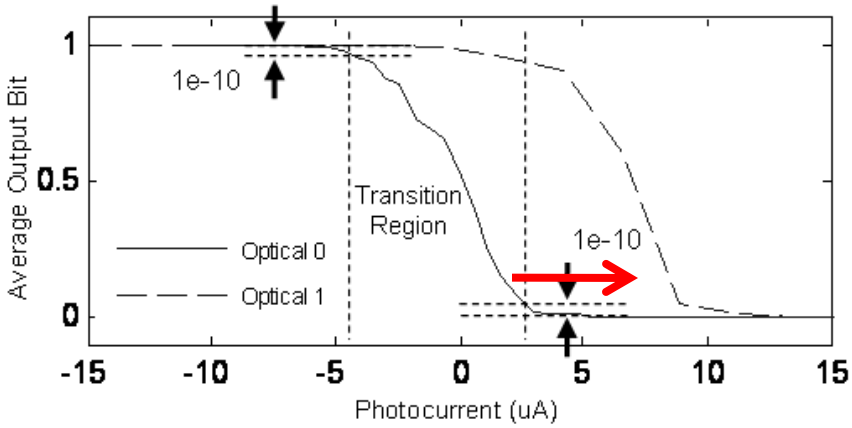


- 10mW required to retune all 8 rings
  - Negligible overhead of tuning circuits (thermal BW < 500kHz)
  - Tuning efficiency 130uW/K (32.4mW/2π) – fully substrate released chips

Low-power current-sensing optical receiver





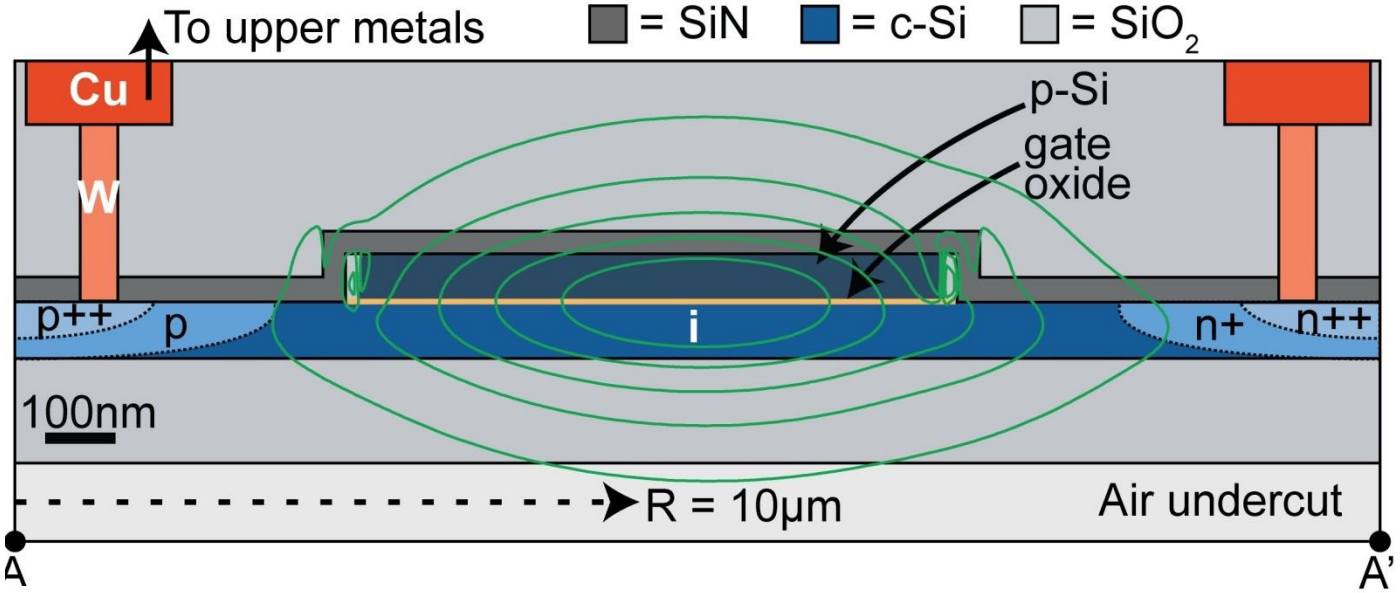


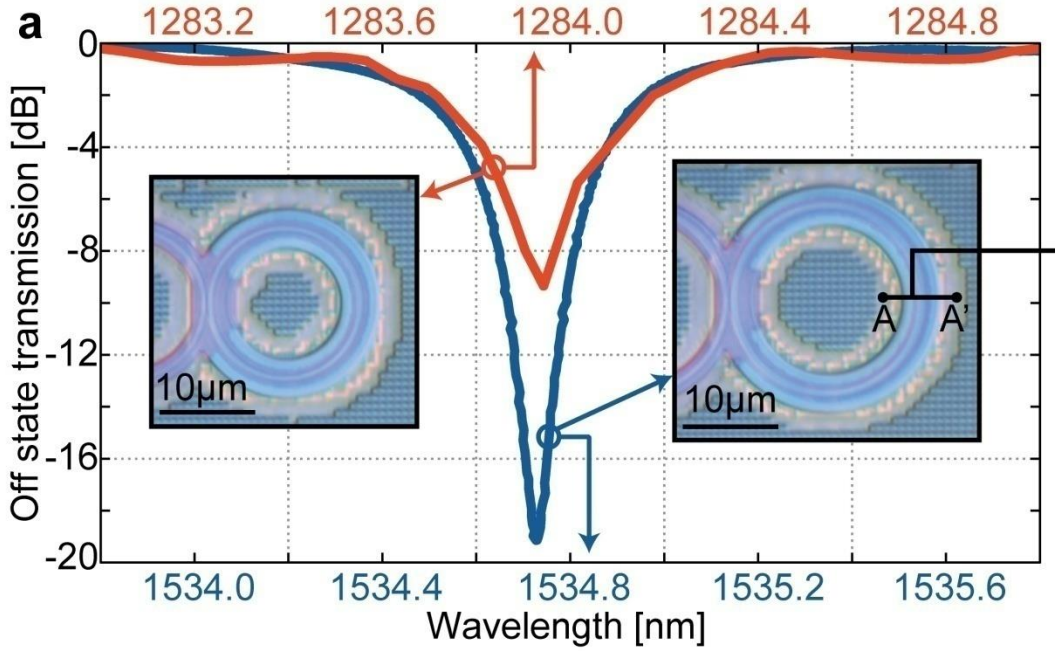
Receiver detects photo current
50fJ/b, uA sensitivities, 3-5Gb/s

Georgas ESSCIRC 2011, JSSC 2012

Optical modulator design

Shainline, Popovic





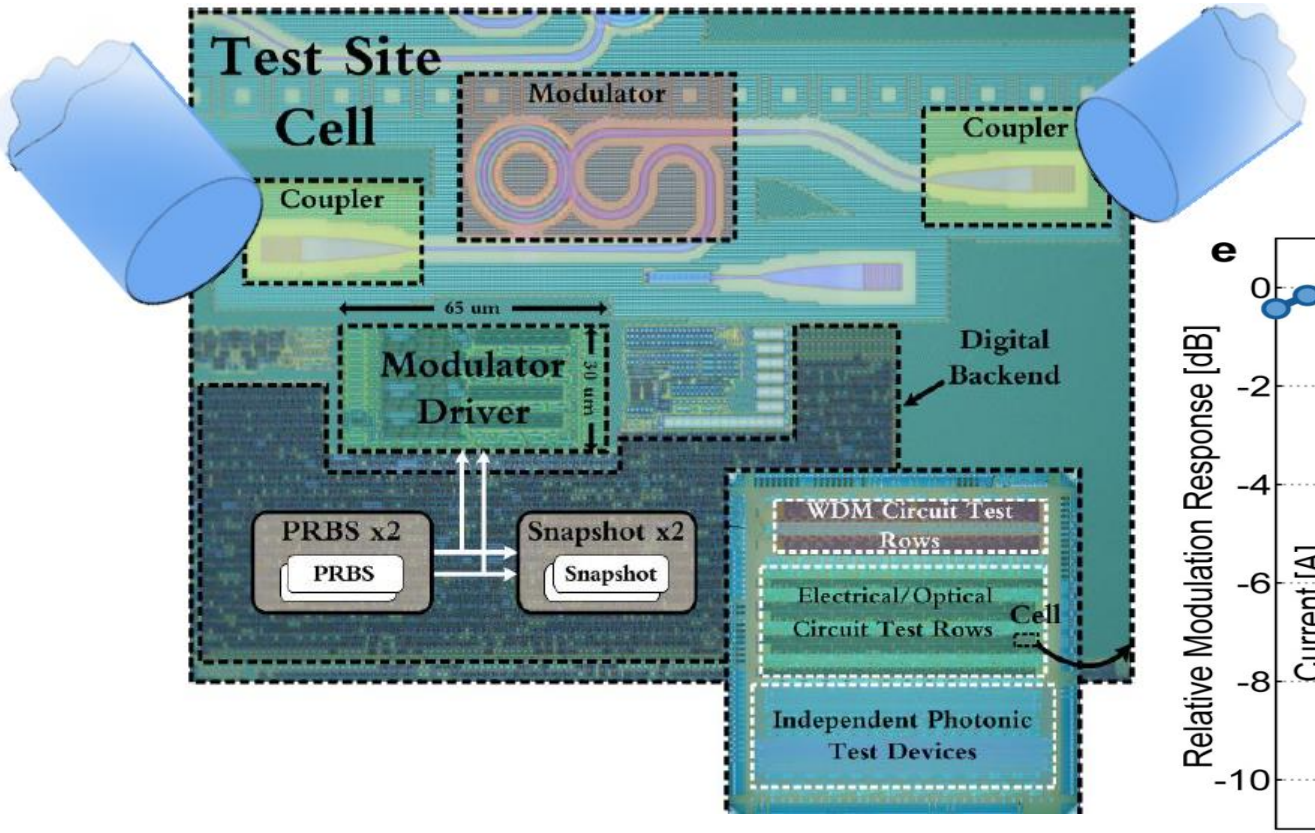
Carrier-injection device at 1550nm

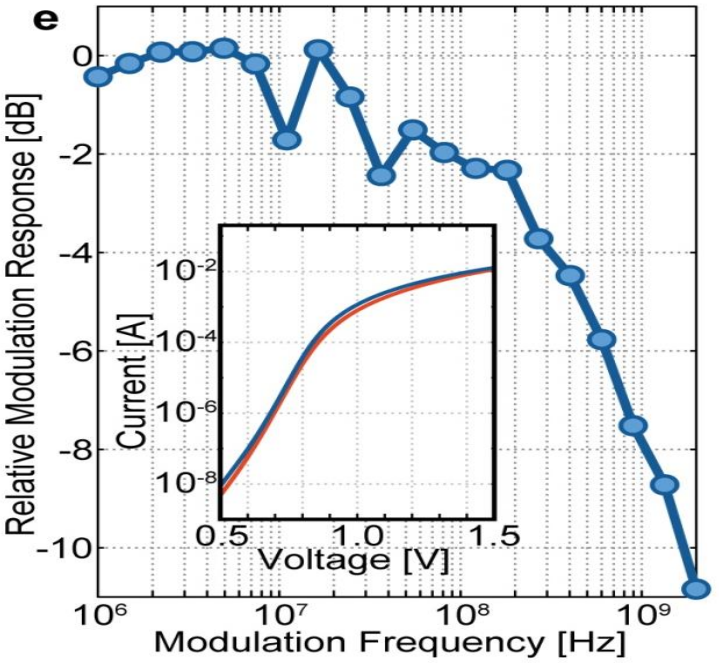
- Extinction ratio 19dB
- 45GHz 3dB optical bw

at 1280nm

- Extinction ratio 9dB
- 60GHz 3dB optical bw

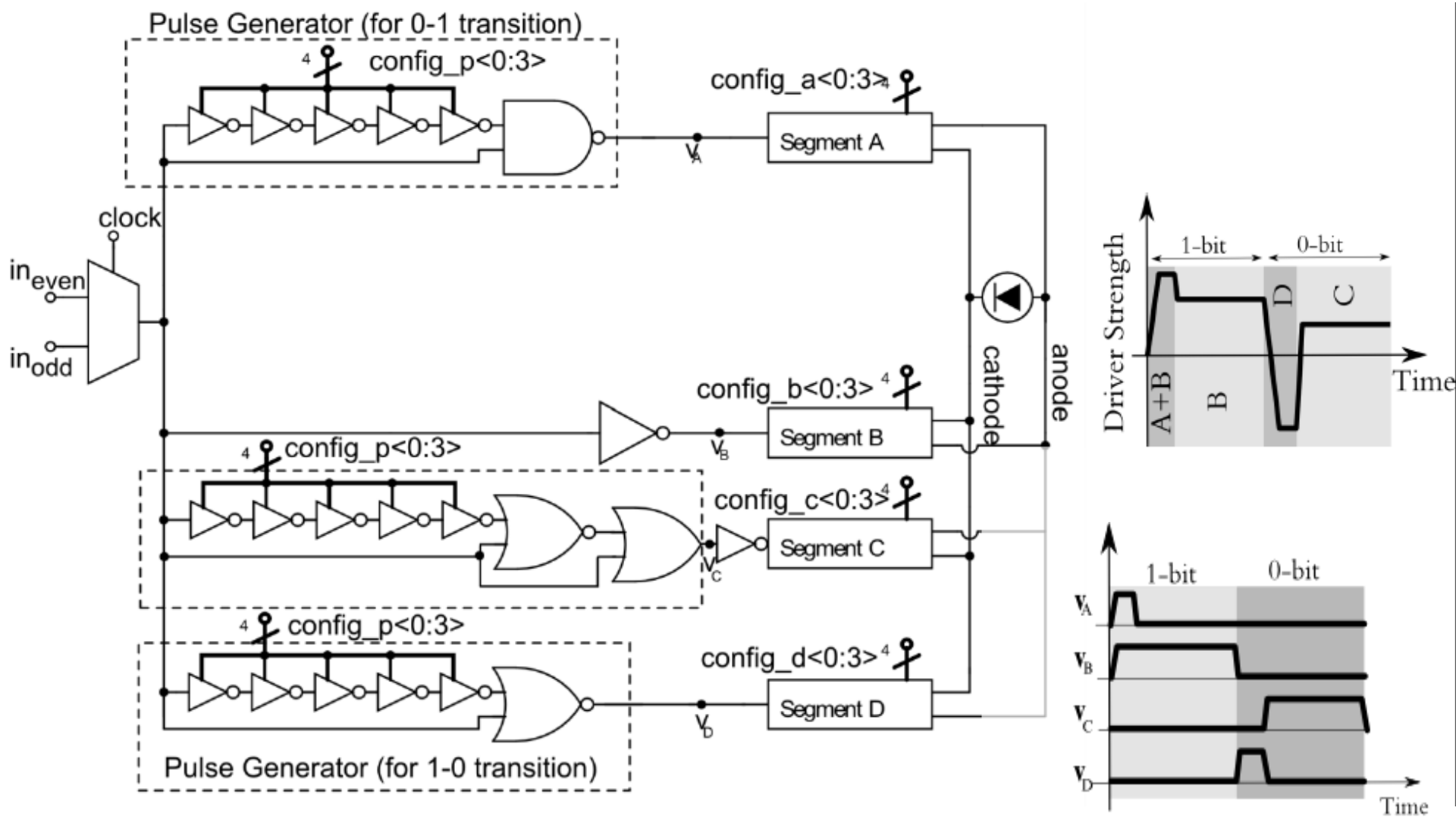
Optical modulator – electrical tests





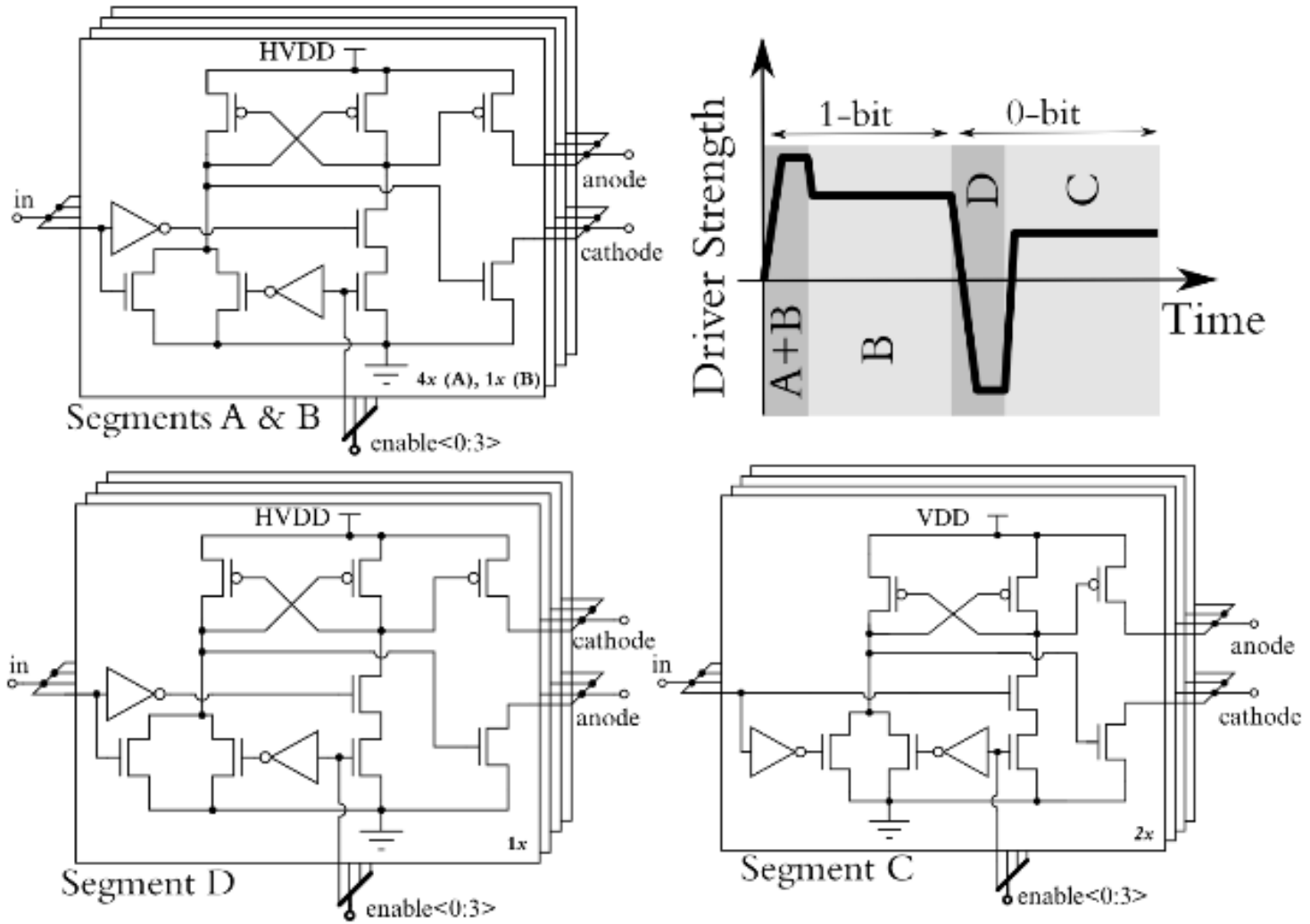
- Carrier-lifetime 2-3ns → **200MHz electrical bandwidth**
  - Diffusion time constant affected by
    - Recombination time
    - Drift conditions

Modulator driver sub-bit pre-emphasis



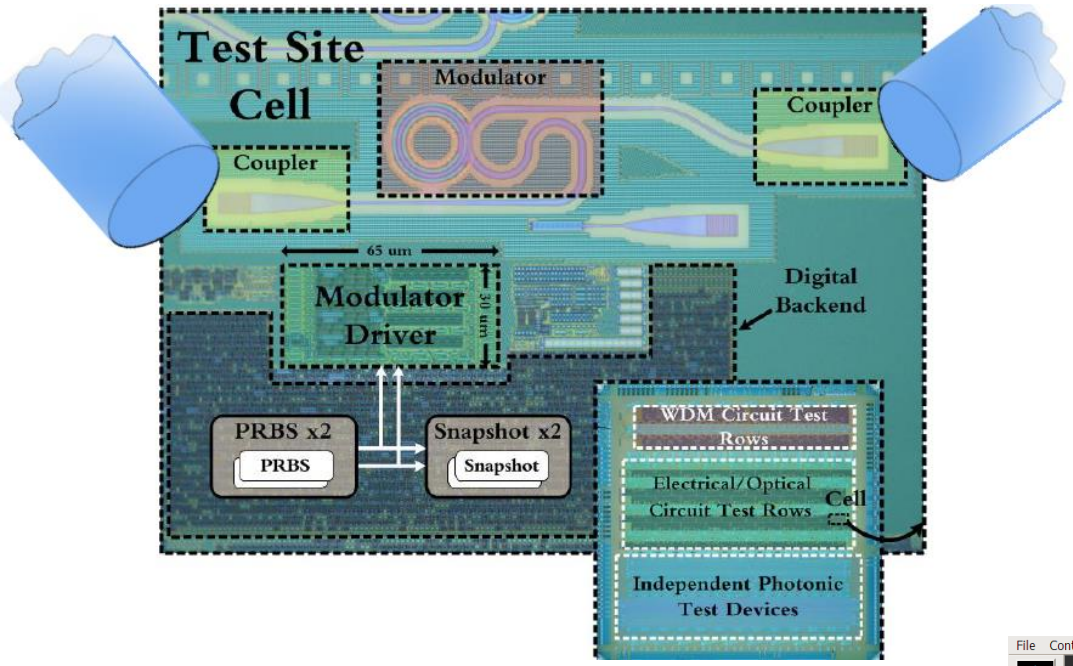
- Partial forward bias at 0-bit key to fast operation

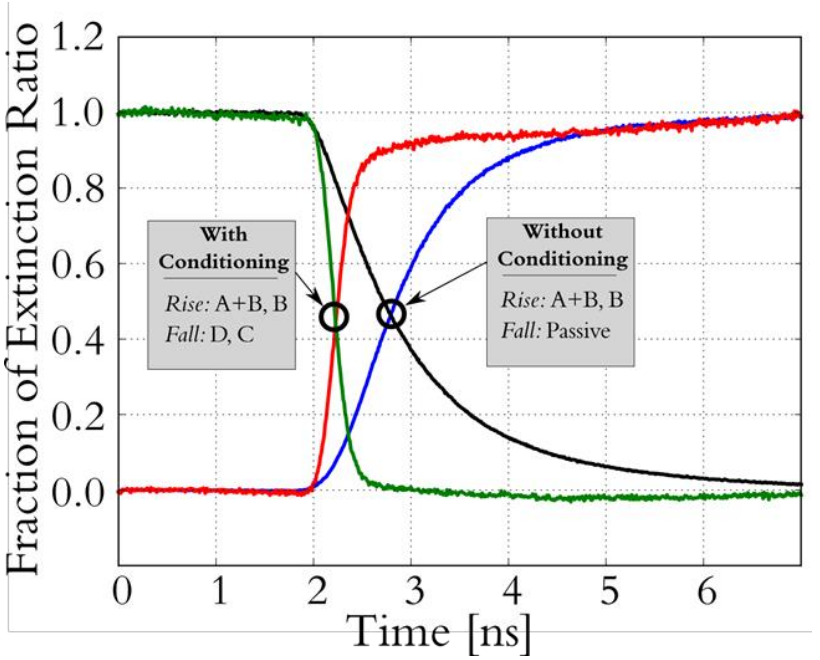
Modulator driver heads



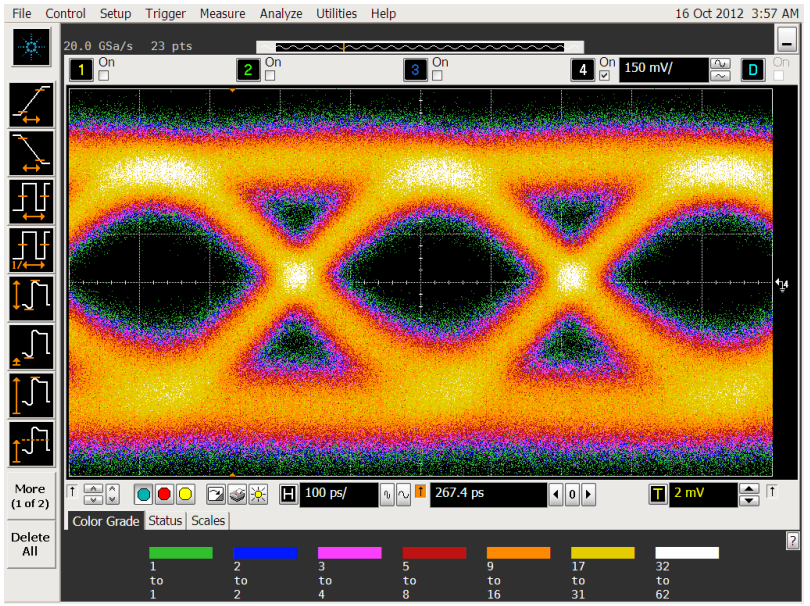
- Split-supply used for sub-bit pre-emphasis
  - Use core and I/O voltage – no regulators

First modulation in 45nm process



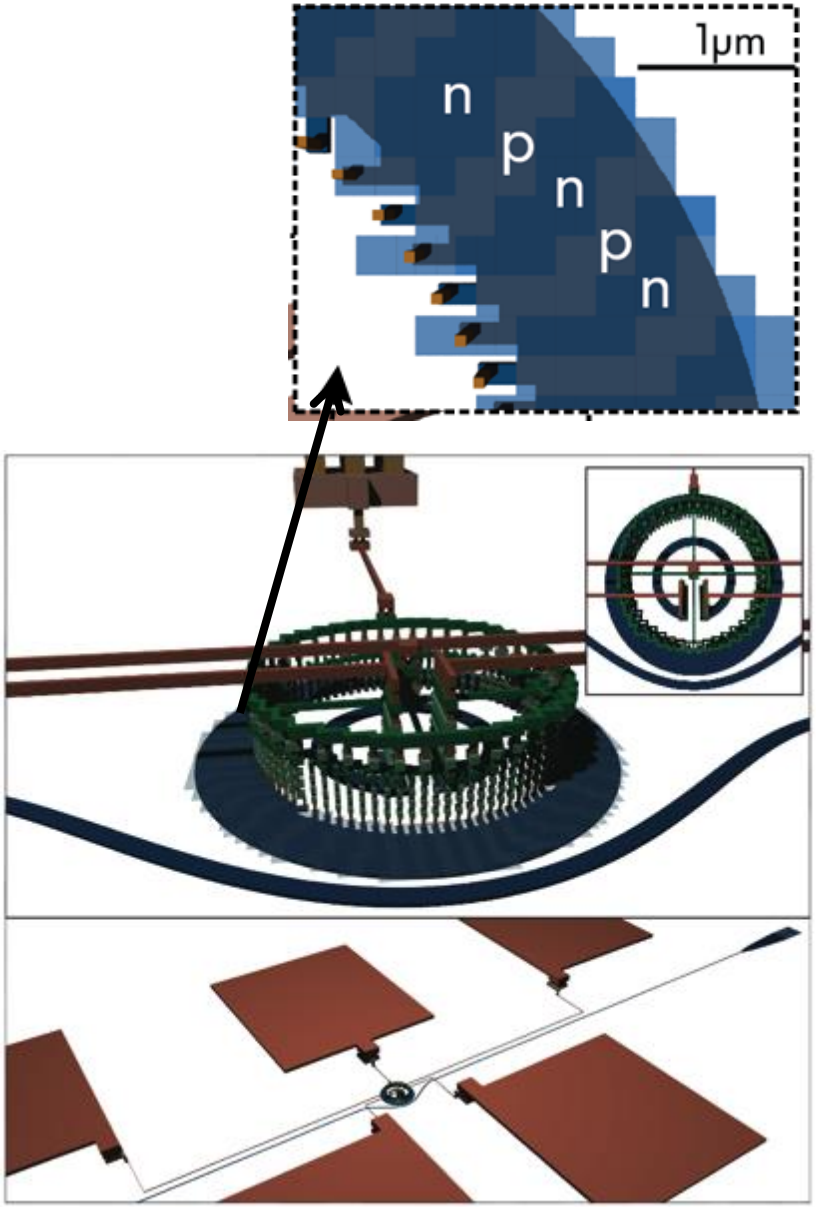


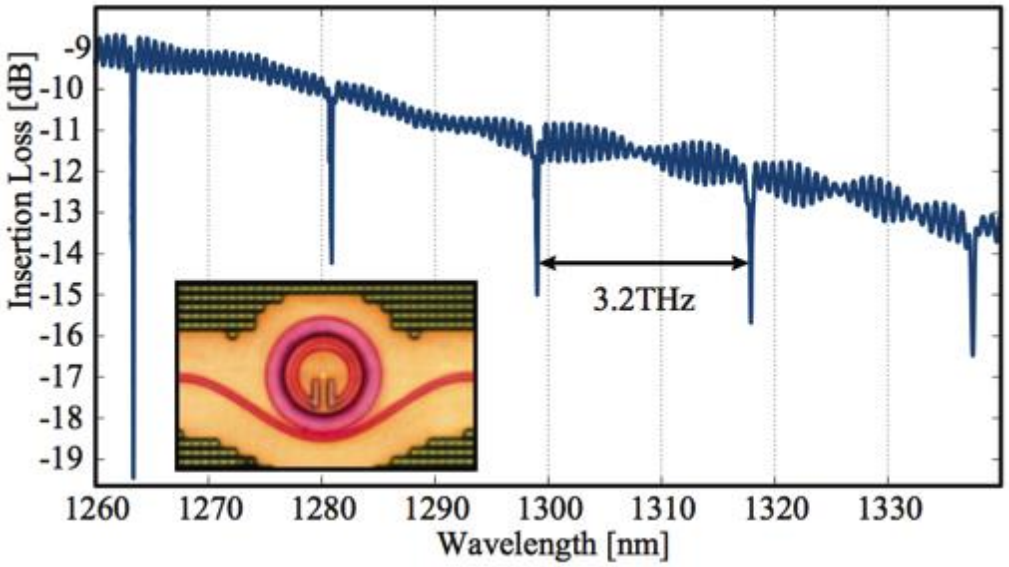
- 2.5Gb/s modulation
- 1.2pJ/bit
- 3dB insertion loss
- 3dB extinction ratio

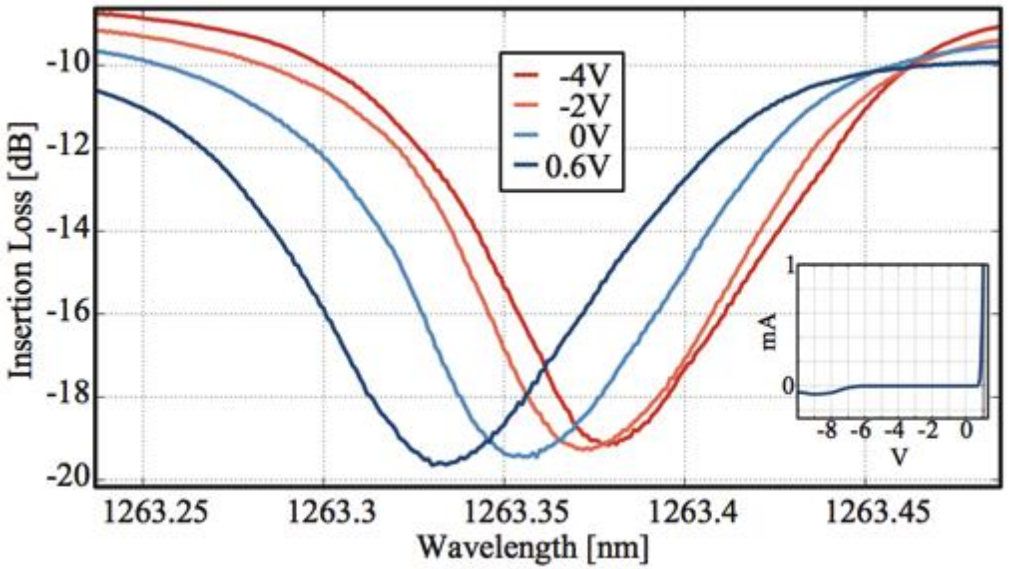


Moss ISSCC 2013

Depletion modulators in 45nm SOI CMOS

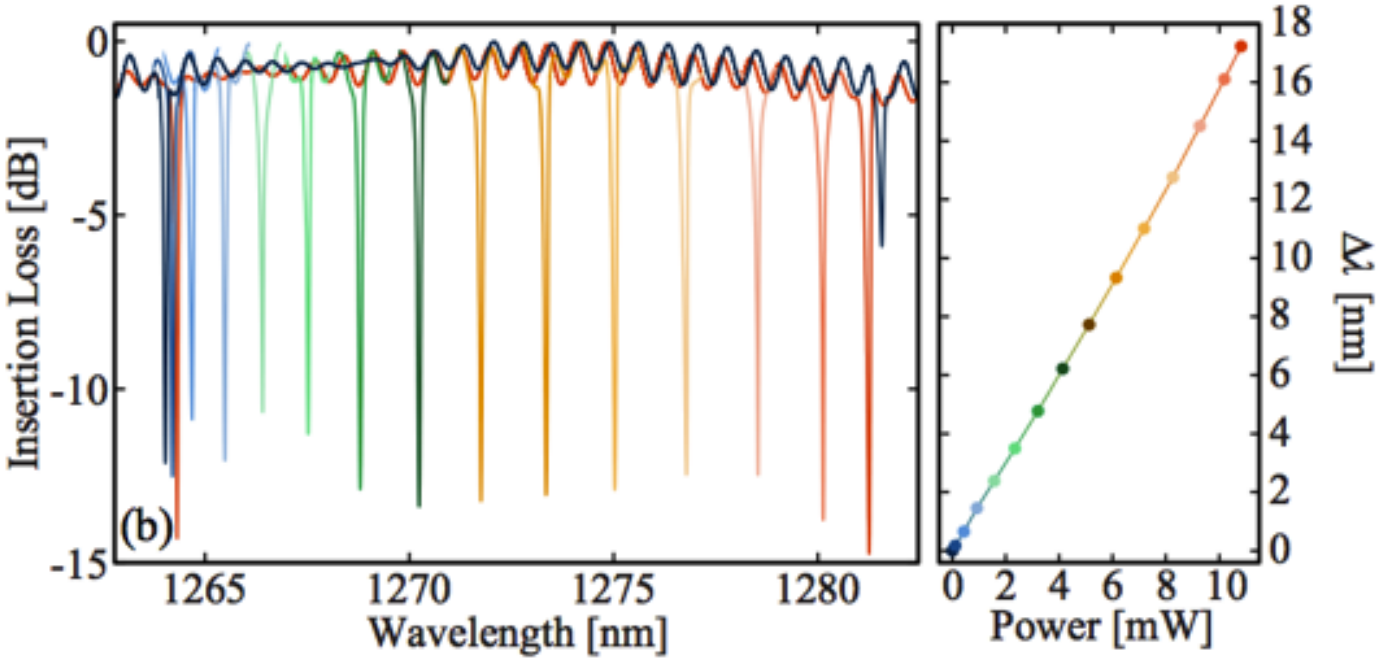




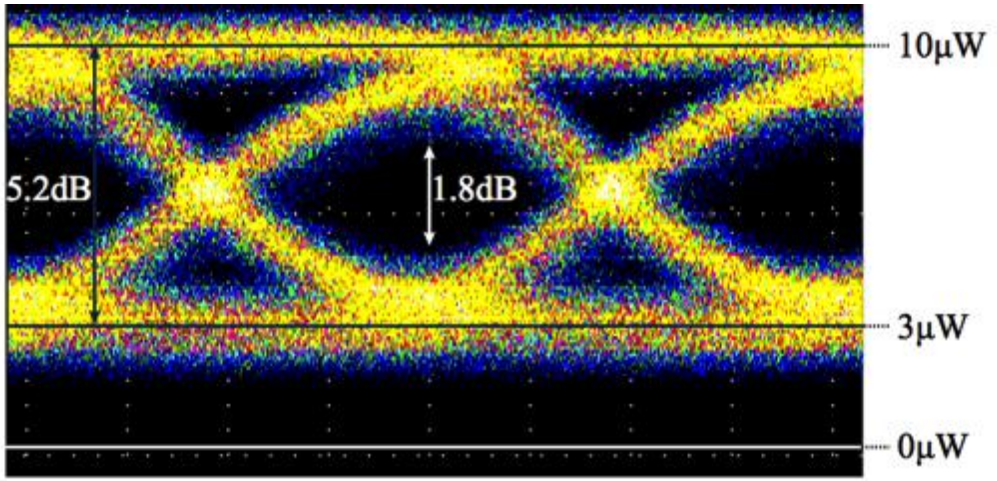


Shainline et al., Optics Letters 2013

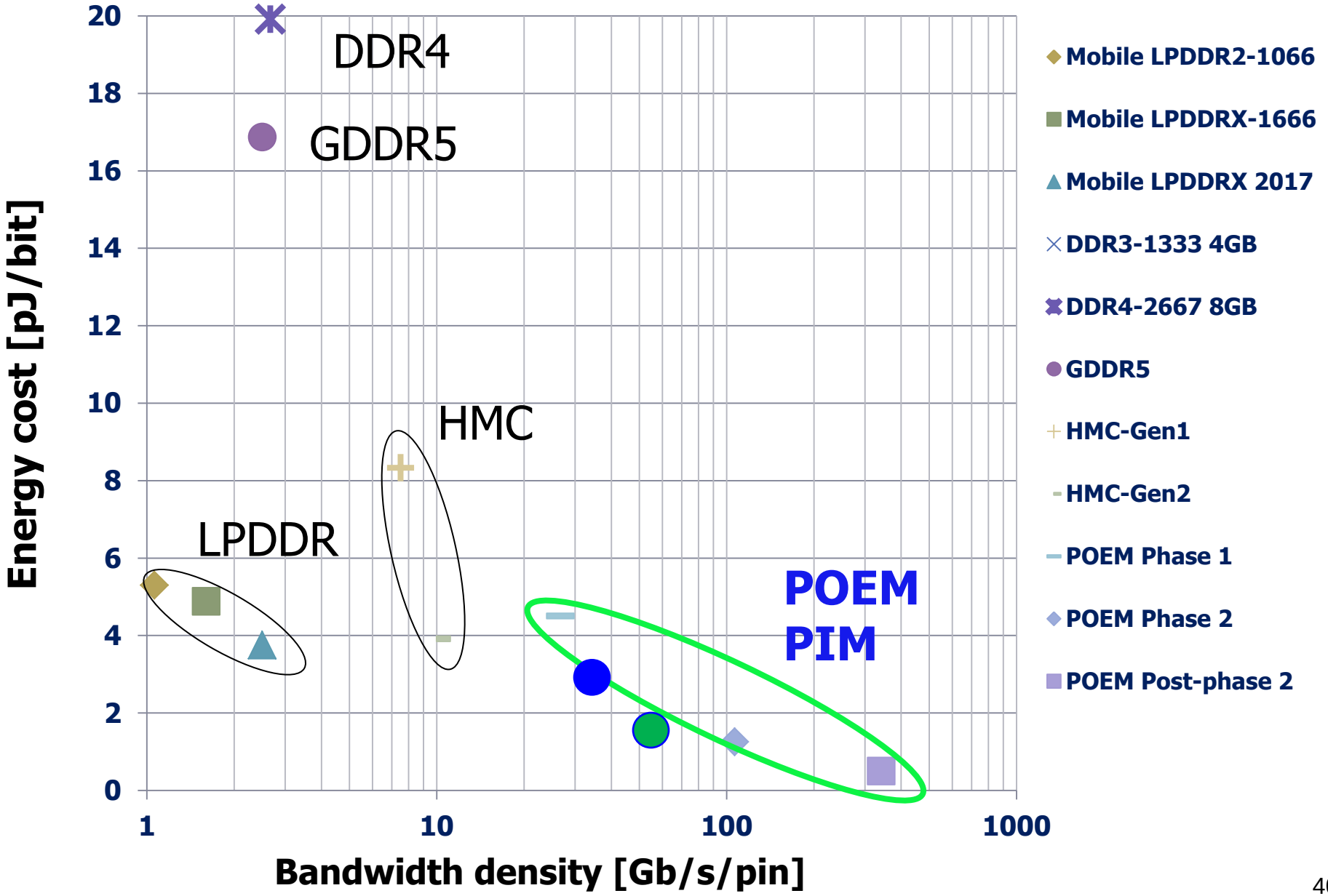
Depletion modulators in 45nm SOI CMOS



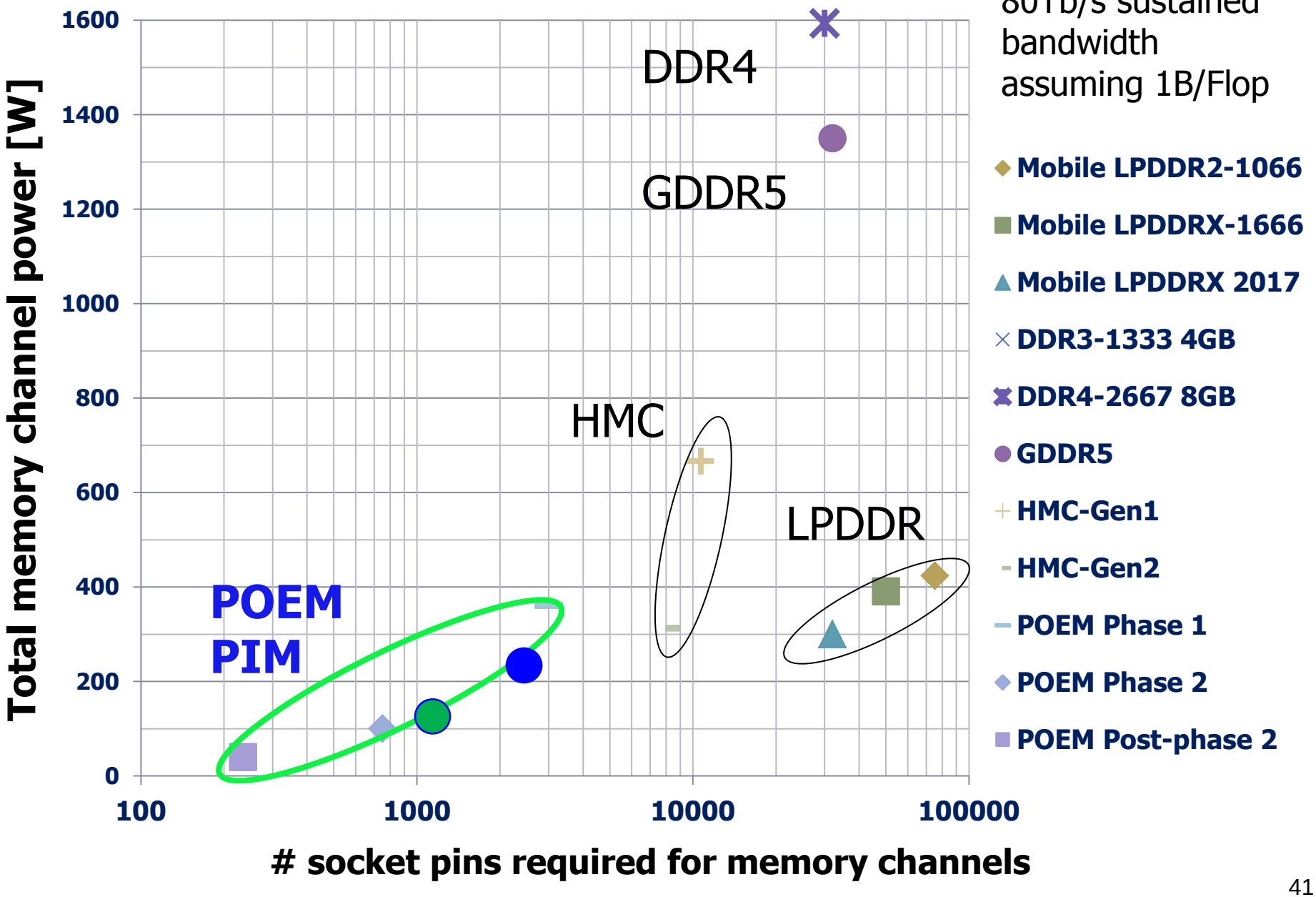
- Modulation:
  - 5 Gbps
  - 5.2dB extinction ratio
- Energy:
  - 55 fJ/bit
  - Tunable across FSR with 400GHz/mW (~2nm/mW)



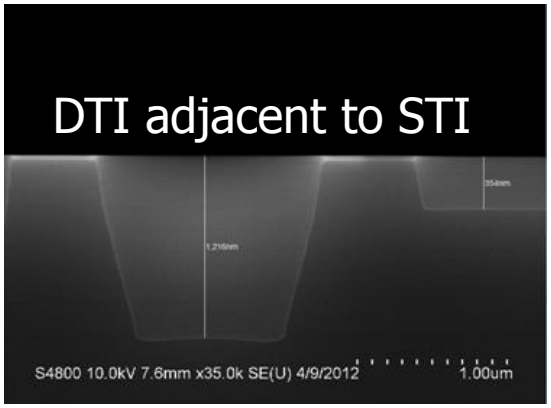
Memory interface scaling problems: Energy-cost and bandwidth density

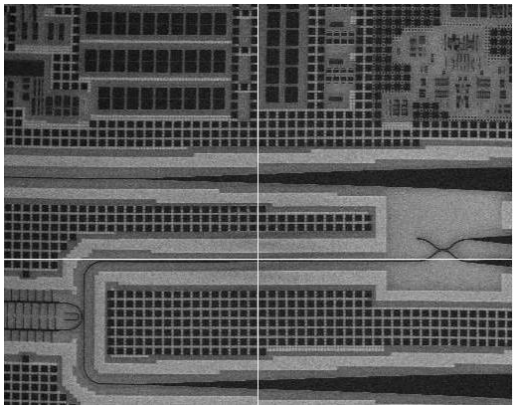


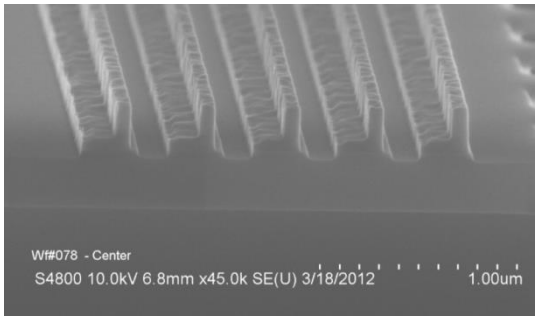
Power and pins required for 10TFlop/s

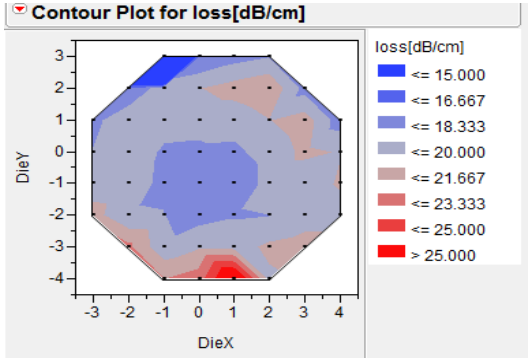


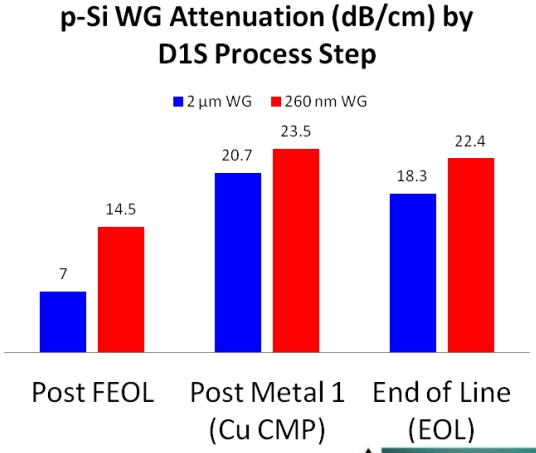
DRAM side: Bulk integration (polySi photonics)







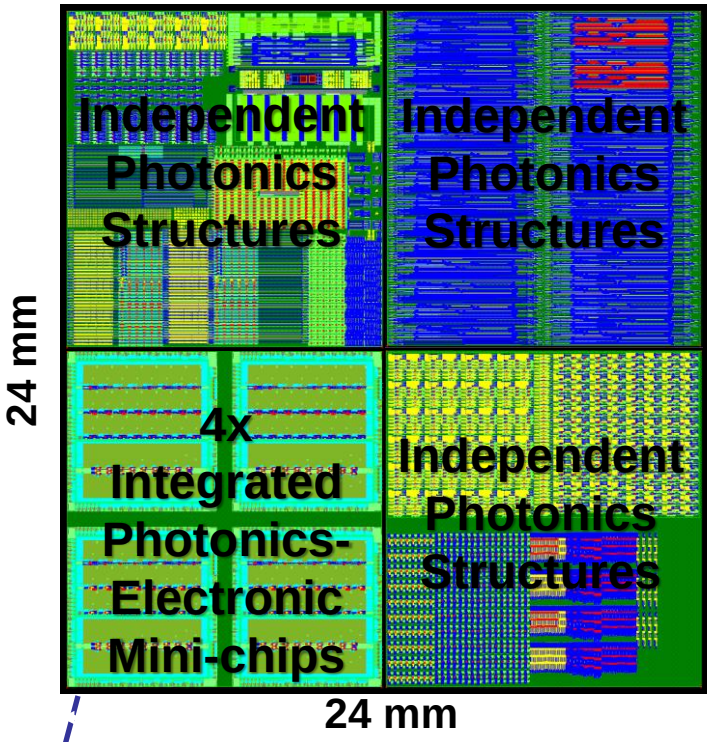


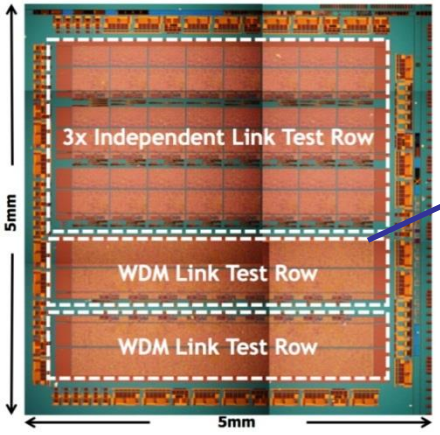


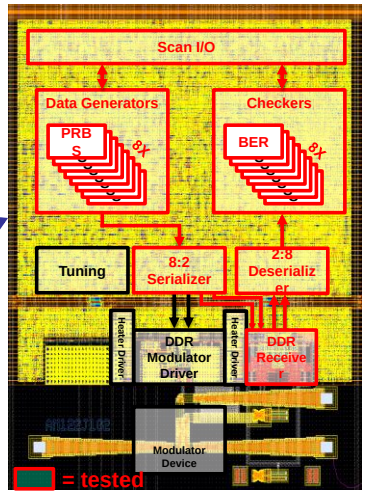
DARPA POEM

Meade et al
Sun et al
OI 2013

Micron Reticle







Summary

- Silicon-photonics can push both critical dimensions
  - Energy-efficiency – monolithic integration
  - Bandwidth Density - dense WDM
- Need to optimize across layers
  - Connect devices to circuits, and links to networks
- Building early technology development platforms
  - Feedback to device and circuit designers
  - Accelerated adoption
- EOS Platform designed for multi-project wafer runs
  - Best end-of-line passives in sub-100nm process (3-4dB/cm loss)
  - sub-100fJ/b transmitters/receivers
  - Record-high tuning efficiency with undercut ~ 25uW/K

Conclusions

- Silicon-photonics – enabler of new capabilities
  - Think “new on-chip inductor” or “new on-chip t-line”
- Potentially revolutionize many applications despite slowdown in CMOS scaling
  - VLSI compute and network infrastructure
  - Wireless comm
  - Imaging and Sensing
- Need process, device, circuit and system-level understanding
- So, jump-in and ride the “new wave”