Inside the PlayStation 5: Solving the Chip Bandwidth Challenge

Ben Moss

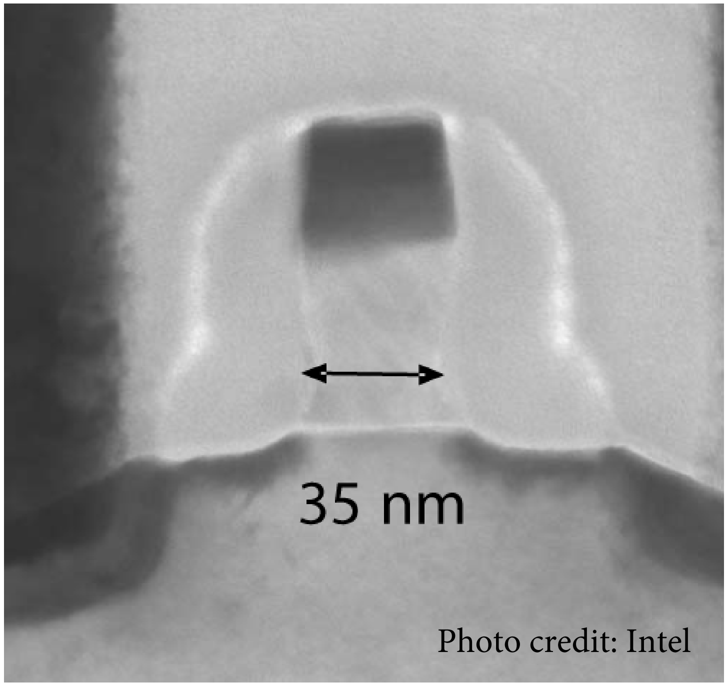
CICS Center for Integrated Circuits and Systems
Massachusetts Institute of Technology

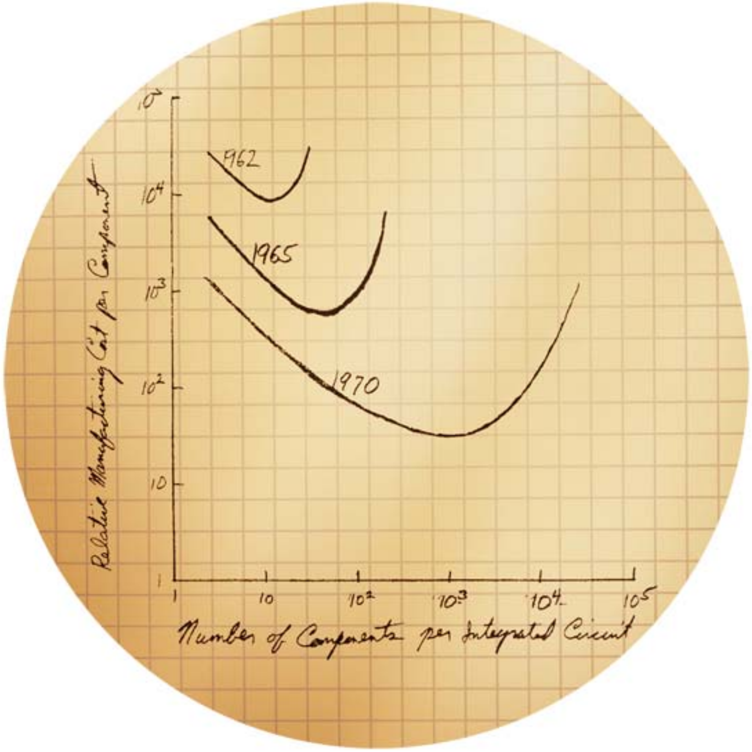
MTL microsystems technology laboratories
massachusetts institute of technology

rle AT MIT

Integrated Systems Group
Massachusetts Institute of Technology

Moore's Law

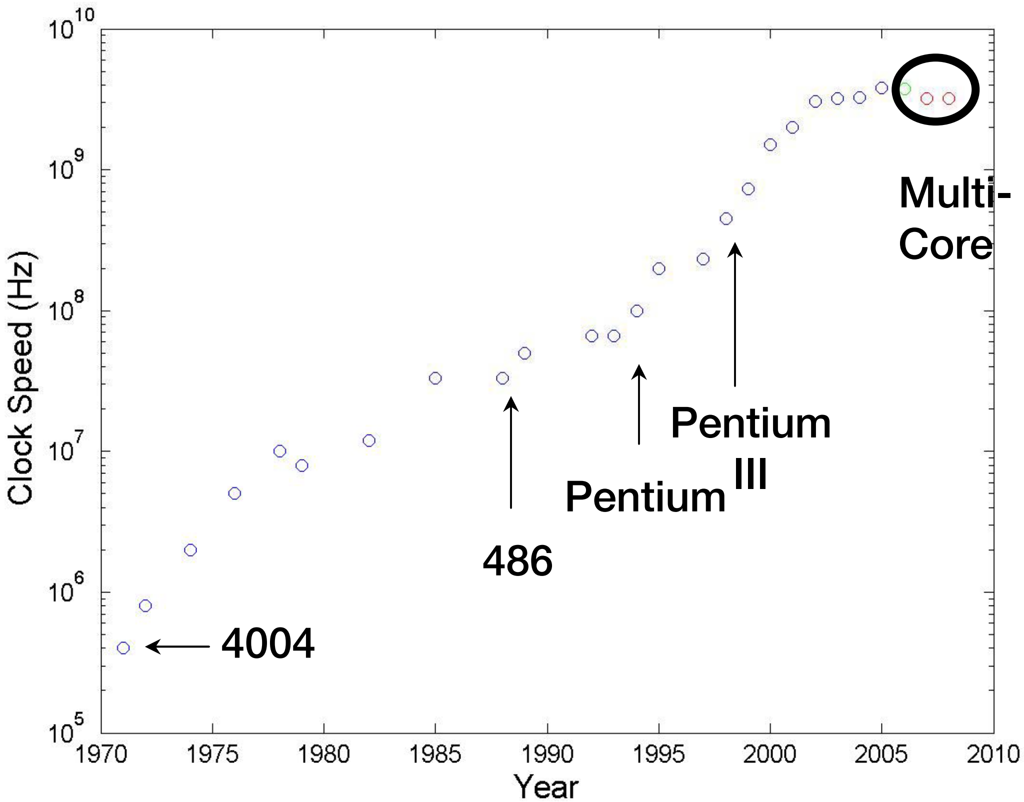




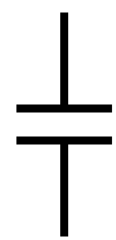
Inexpensive transistor density
doubles roughly every two years

Processor Parallelism

Intel Desktop Processors: 1970 - Now



- Power Constrained
- No power density scaling
- No battery energy scaling



$$P = fCV^2$$

Inside the PlayStation 3: IBM Cell

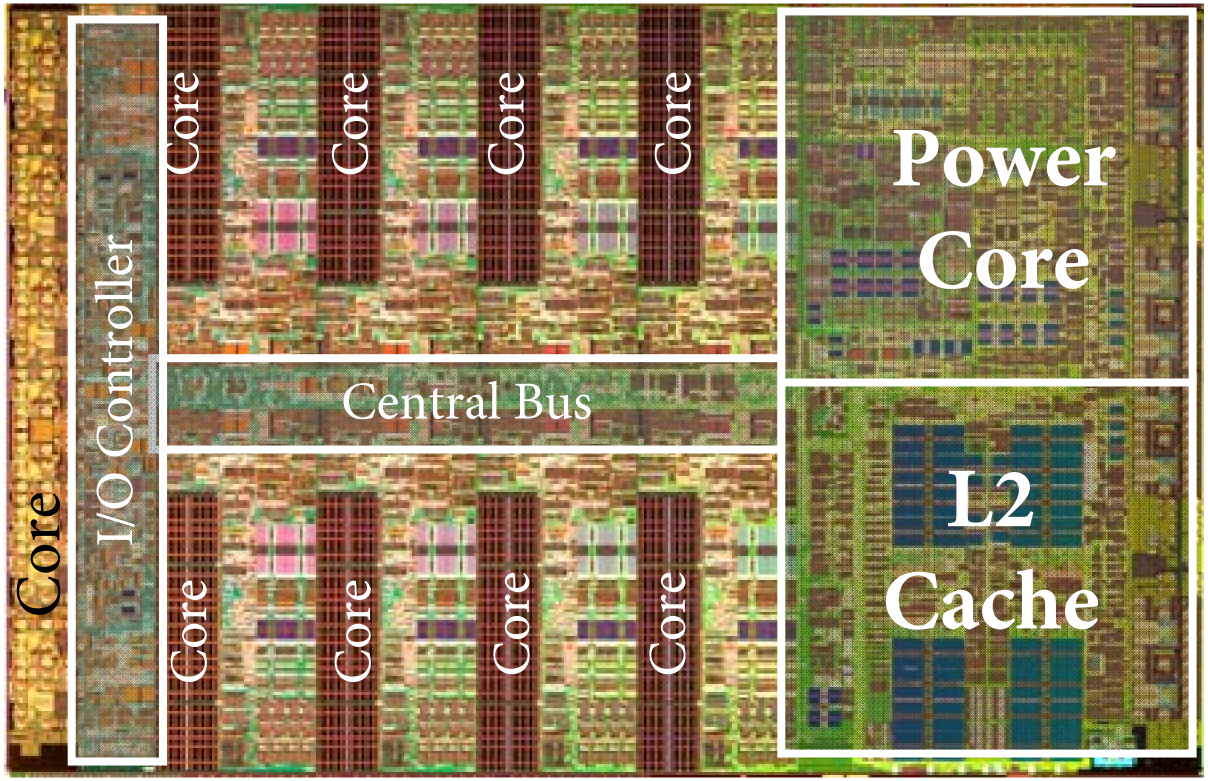
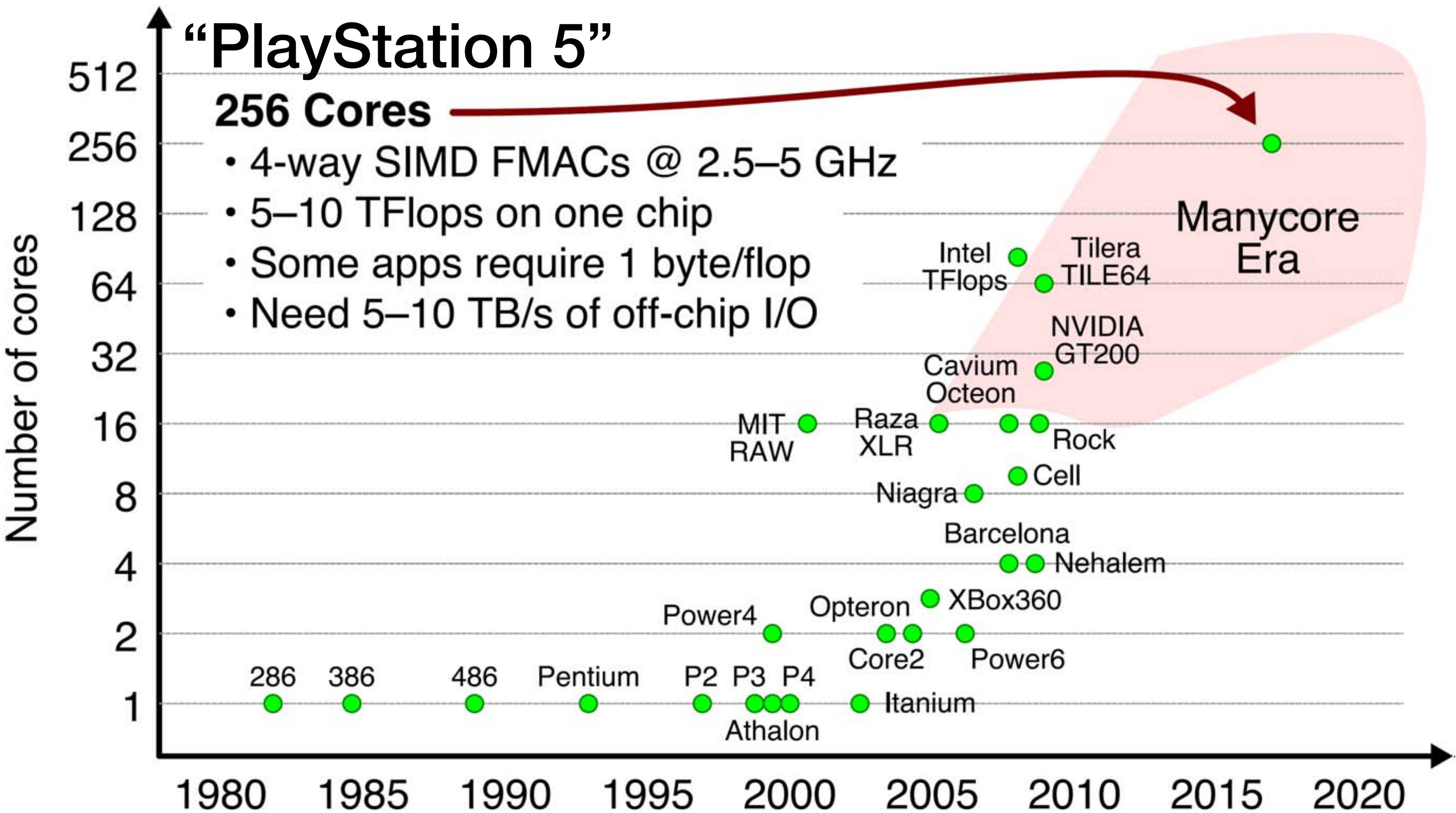


Photo Source: IBM

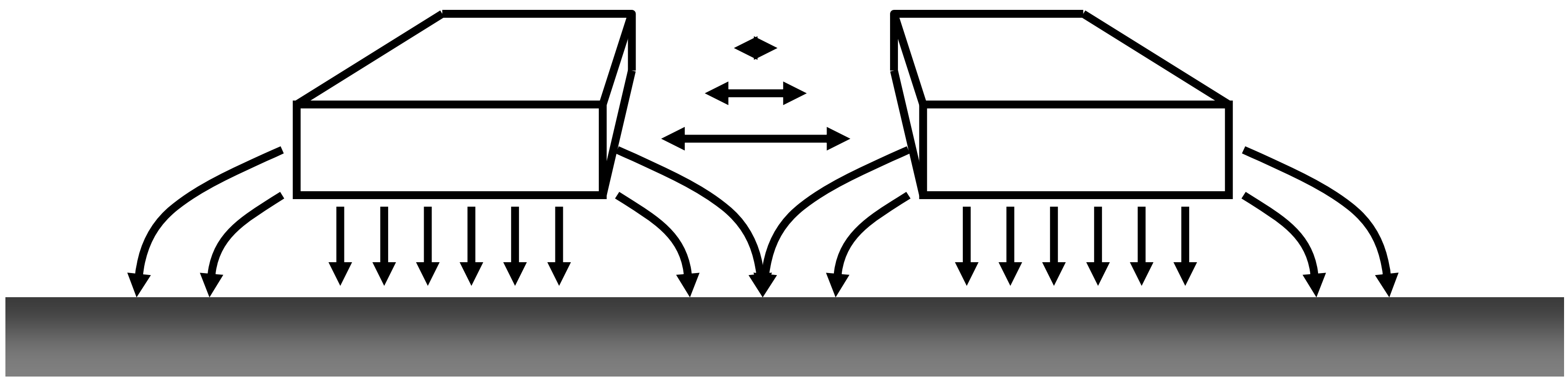
- How do we make the PS5? Add more cores!

Manycore System Bandwidth Needs

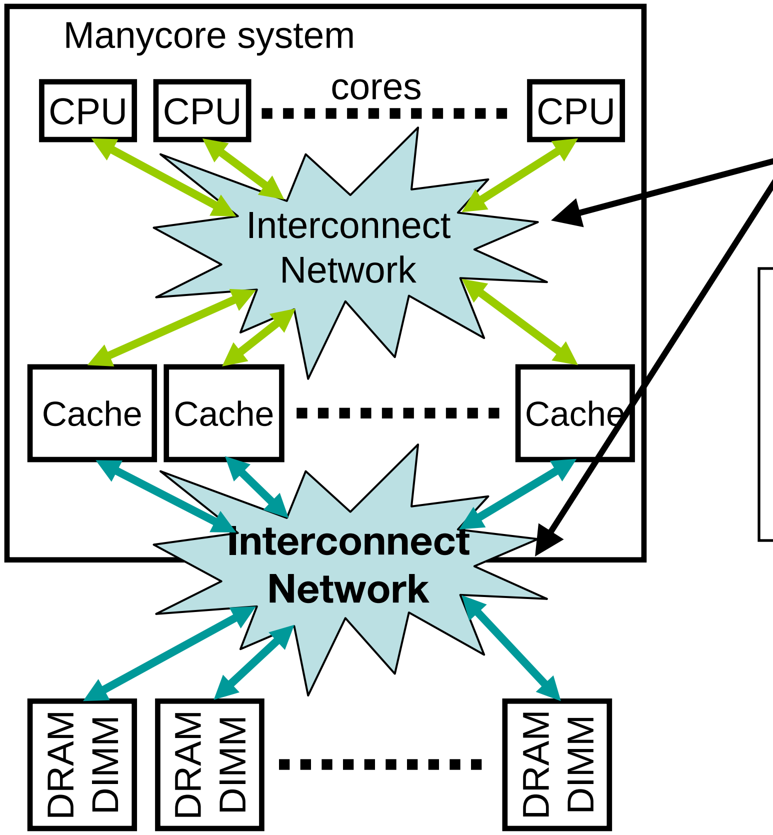


Interconnect Scaling: Not Good

- Wire gets smaller → R increases
- Wire gets smaller →fringe cap dominates
- Dense wires → high crosstalk
- Overall, RC time constant is not improving



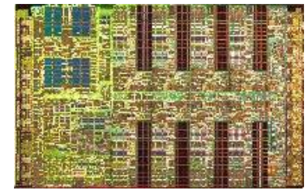
Manycore Bandwidth Bottlenecks



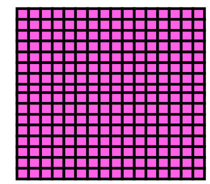
Bottlenecks due to energy and bandwidth density limitations

Constraints:

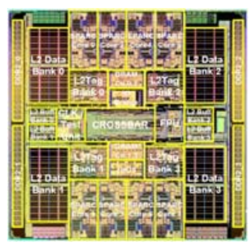
- Area (Bandwidth Density)
- Energy Efficiency
- Compatibility with Bulk-CMOS



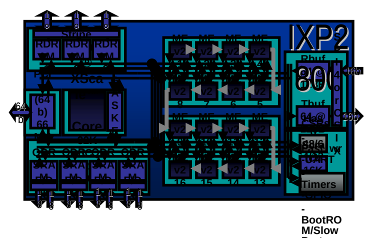
IBM Cell
1 GPP (2 threads)
8 ASPs



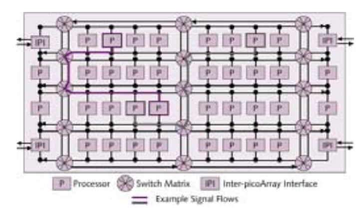
Cisco CSR-1
188 Tensilica GPPs



Sun Niagara
8 GPP cores
(32 threads)



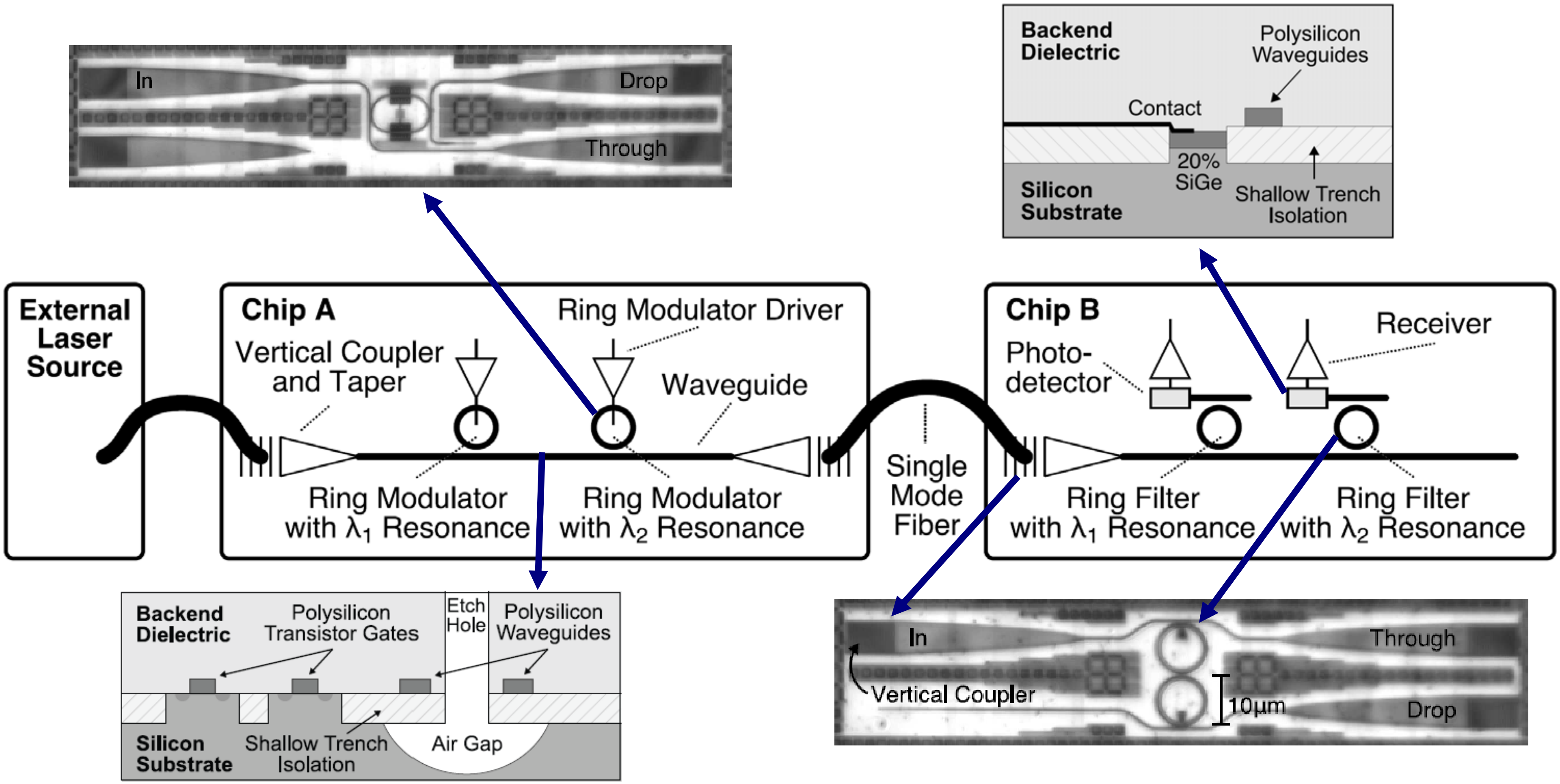
Intel Network Processor
1 GPP Core
16 ASPs (128 threads)



Picochip DSP
1 GPP core
248 ASPs

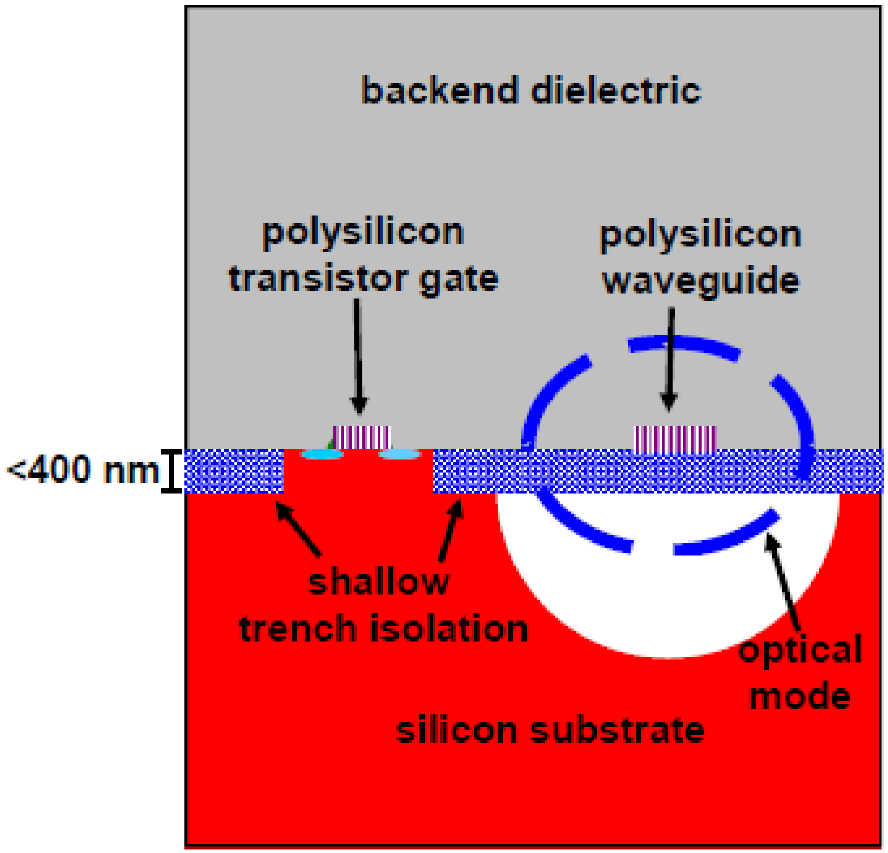
- Conceived in the 1980's
- Recently became feasible
- Must be compatible with Bulk-CMOS
- Need a fair electrical comparison

An Optical Link



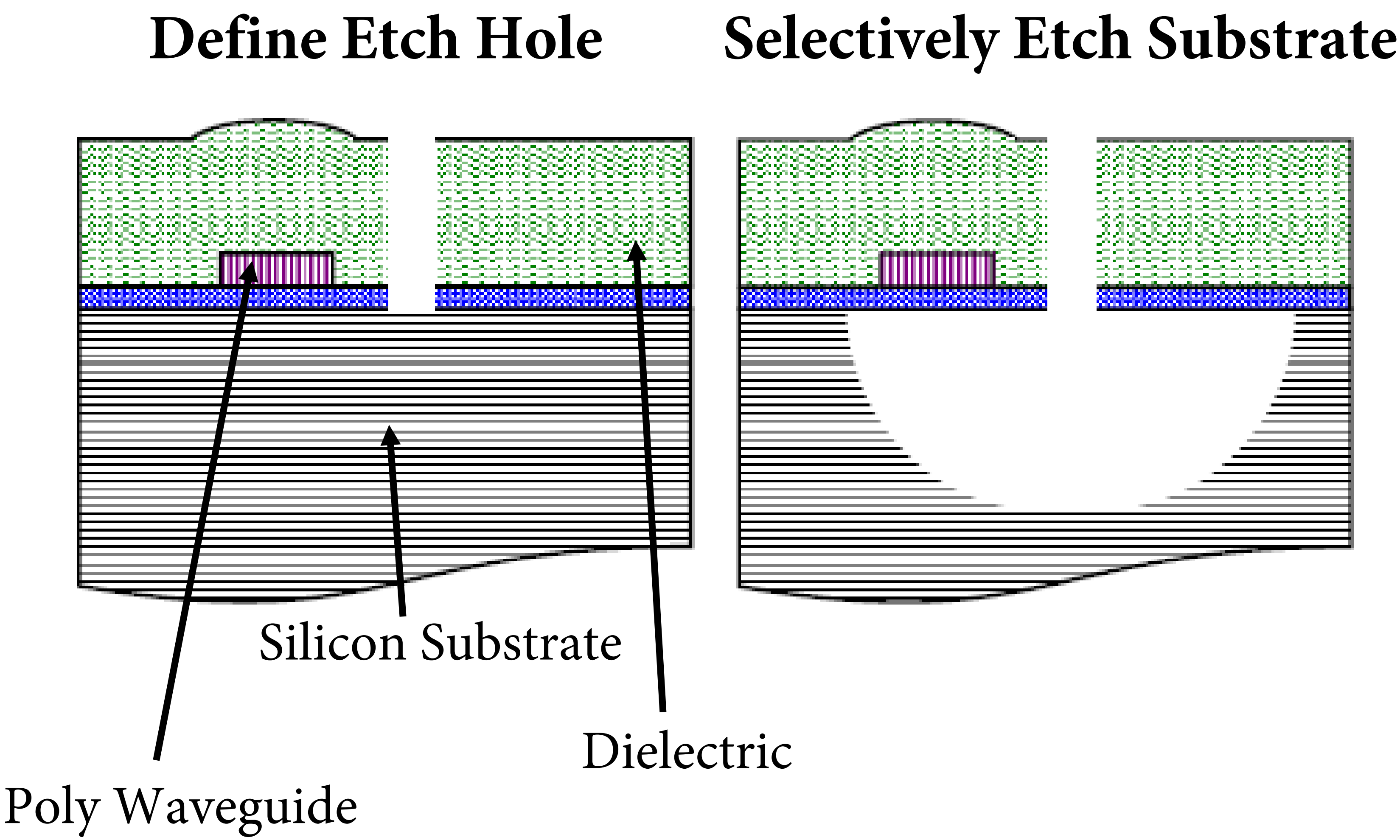
65 nm bulk CMOS chip designed to test various optical devices

Optical Devices: Waveguides

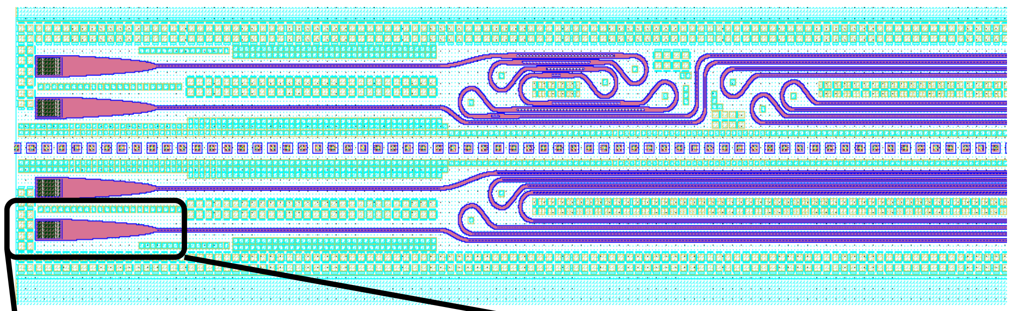


- Basic Building Block
- Poly (transistor gate) layer
- ~0.5um wide, 0.1um tall
- Must be undercut
- Bi-directional
- Can turn, bend, cross!
- WDM

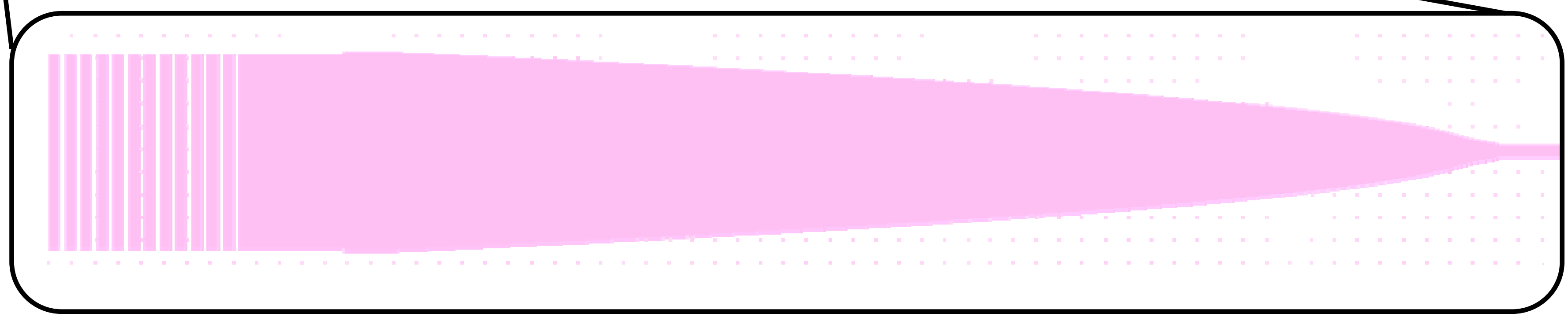
Post-Processing the Waveguide



Optical Devices: Vertical Couplers

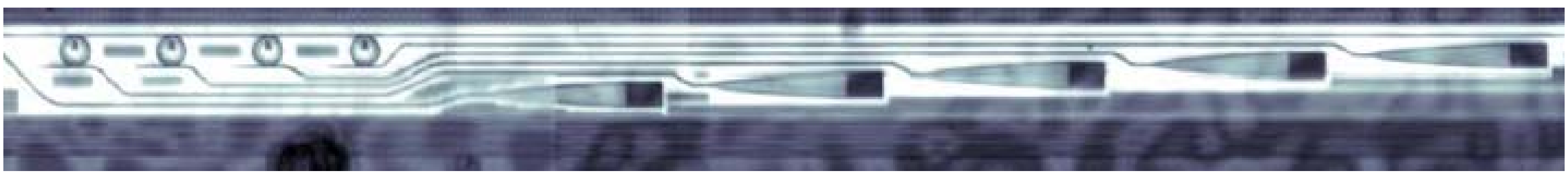


- Vital component
- Sensitive to fiber position
- Also made of poly
- Can enable 3D stacking

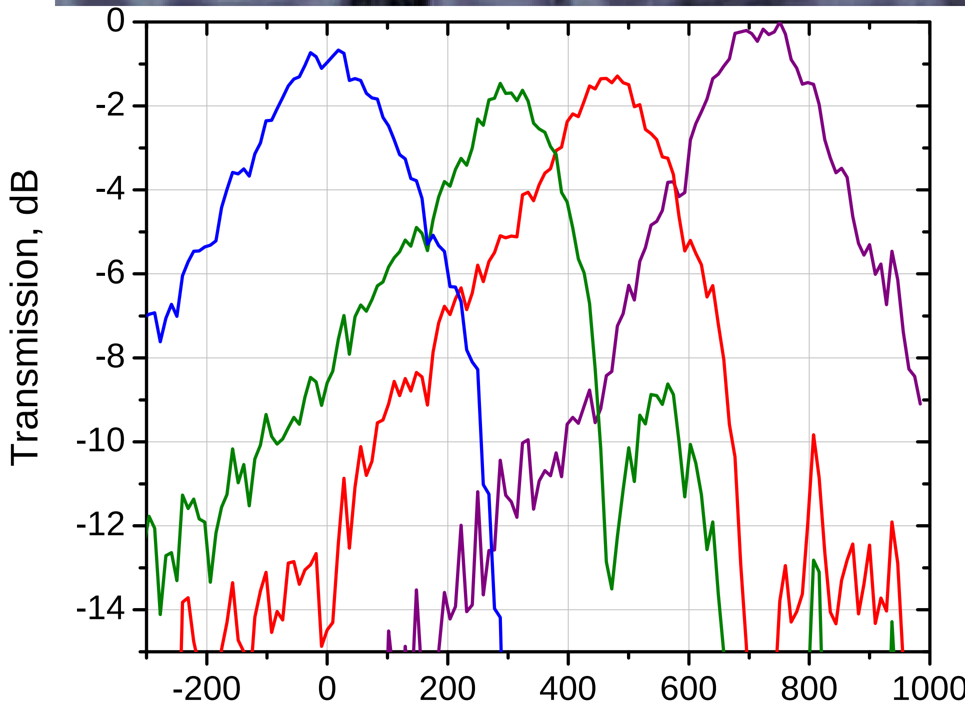


Optical Devices: Ring Filters

- Drops one λ onto a different waveguide
- Geometry is very sensitive
- 1 THz/nm sensitivity on thickness
- 30 GHz/nm sensitivity on width
- Thermally sensitive



Optical Devices: Ring Filter Bank



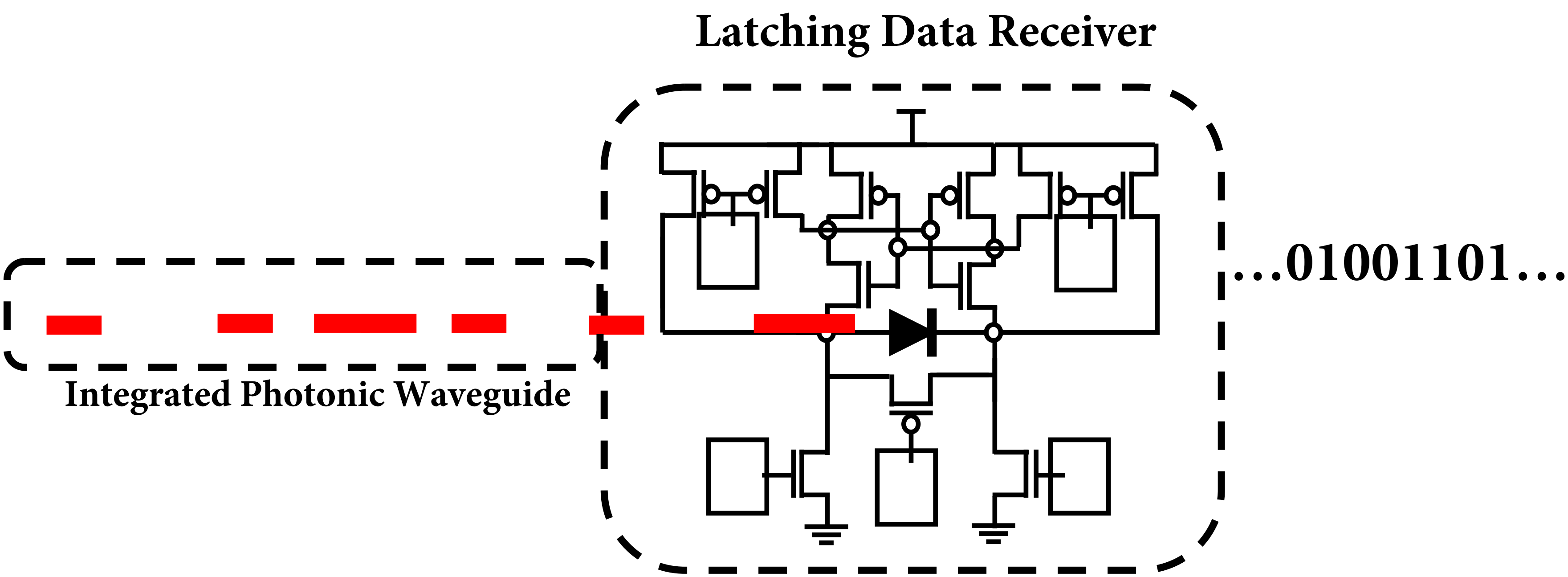
- FSR is ~16nm (2.7THz)
- The scans are 1268nm - 1276nm
- 5nm ring radius step
- Sensitivity
  - 1THz/nm poly
  - 30GHz/nm poly W

Orcutt, J. S., A. Khilo, M. A. Popovic, C. W. Holzwarth, B. Moss, H. Li, M. S. Dahlem, T. D. Bonifield, F. X. Kärtner, E. P. Ippen, J. L. Hoyt, R. J. Ram, and V. Stojanovic, "Demonstration of an Electronic Photonic Integrated Circuit in a Commercial Scaled Bulk CMOS Process," *Optical Society of America - CLEO/QELS Conference*, May 2008.

First 100 GHz spaced bank in sub-100nm, first in bulk CMOS, first in poly Si

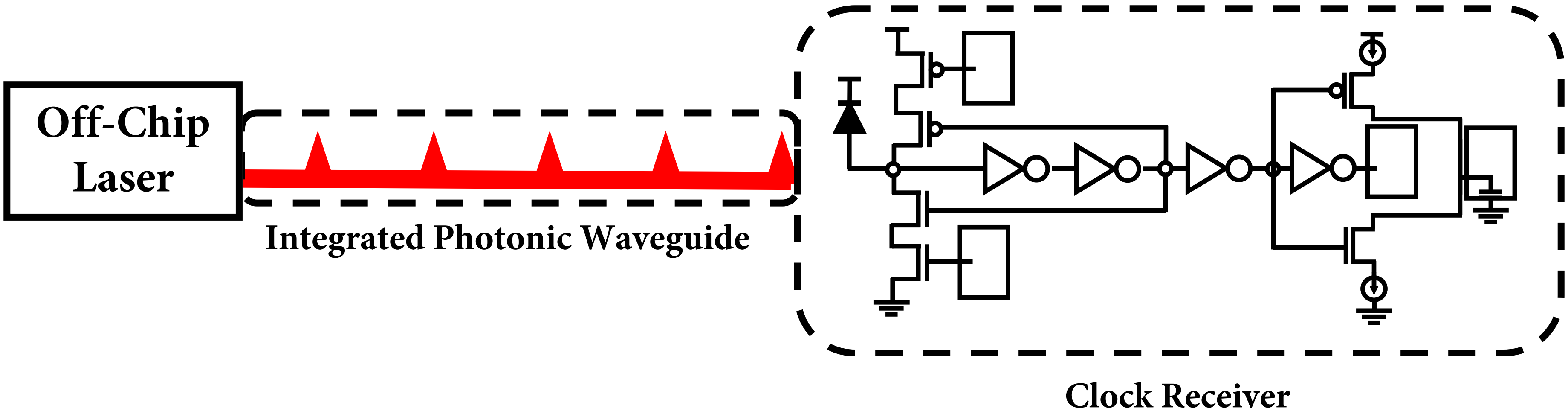
- Enables 60-120 wavelengths/waveguide (>200 Gb/s/um data rate density)

Optical Devices - Photodiode



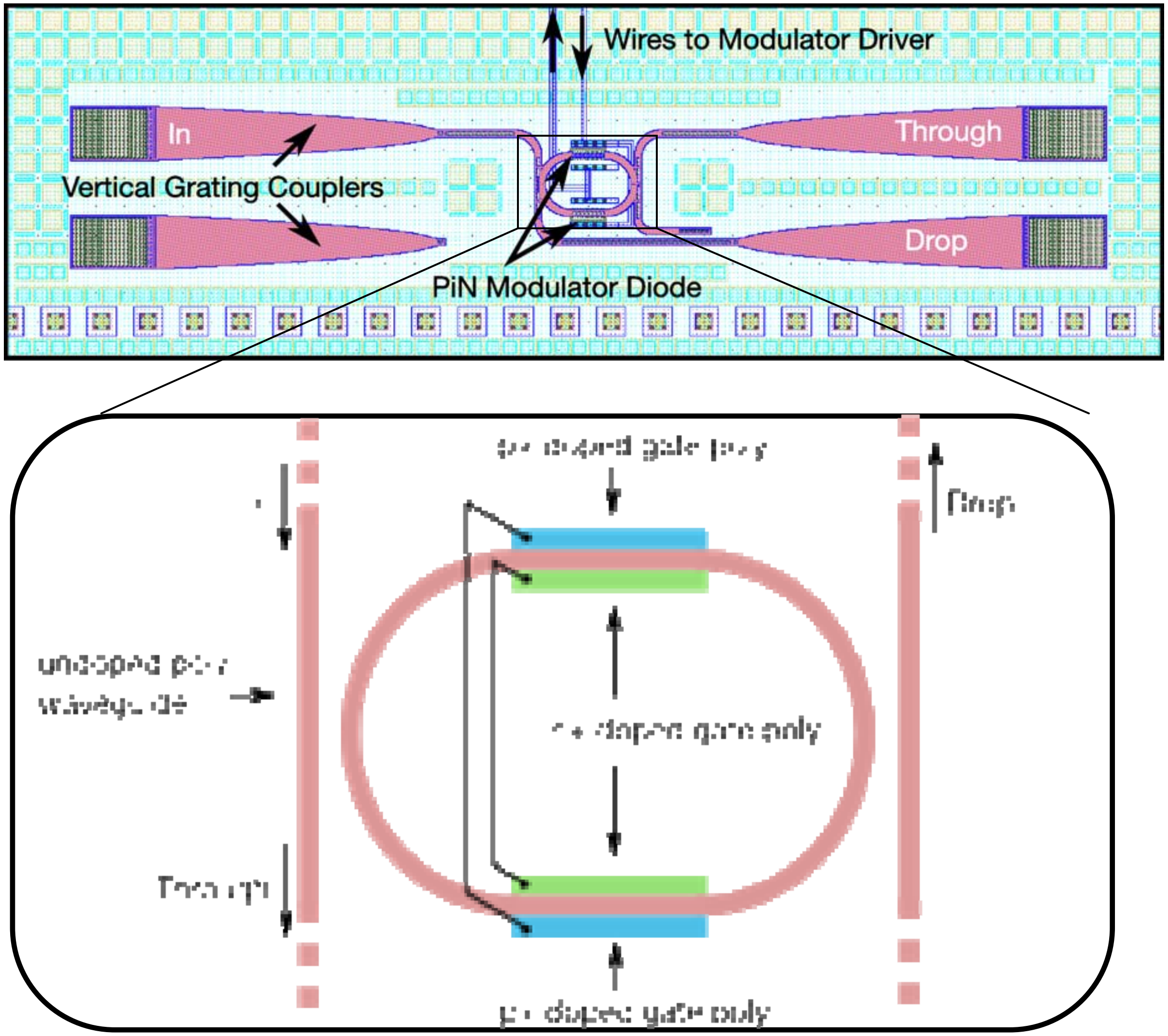
The current signal from the photodiode goes to a receiver

Optical Clocking



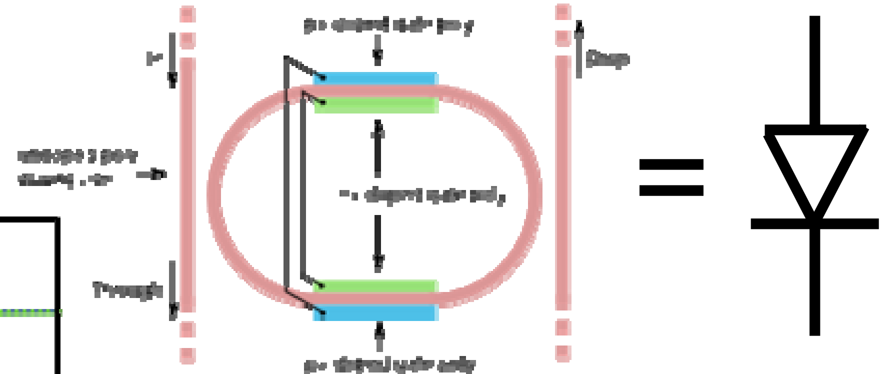
Optical clocking can reduce clock skew and increase energy efficiency

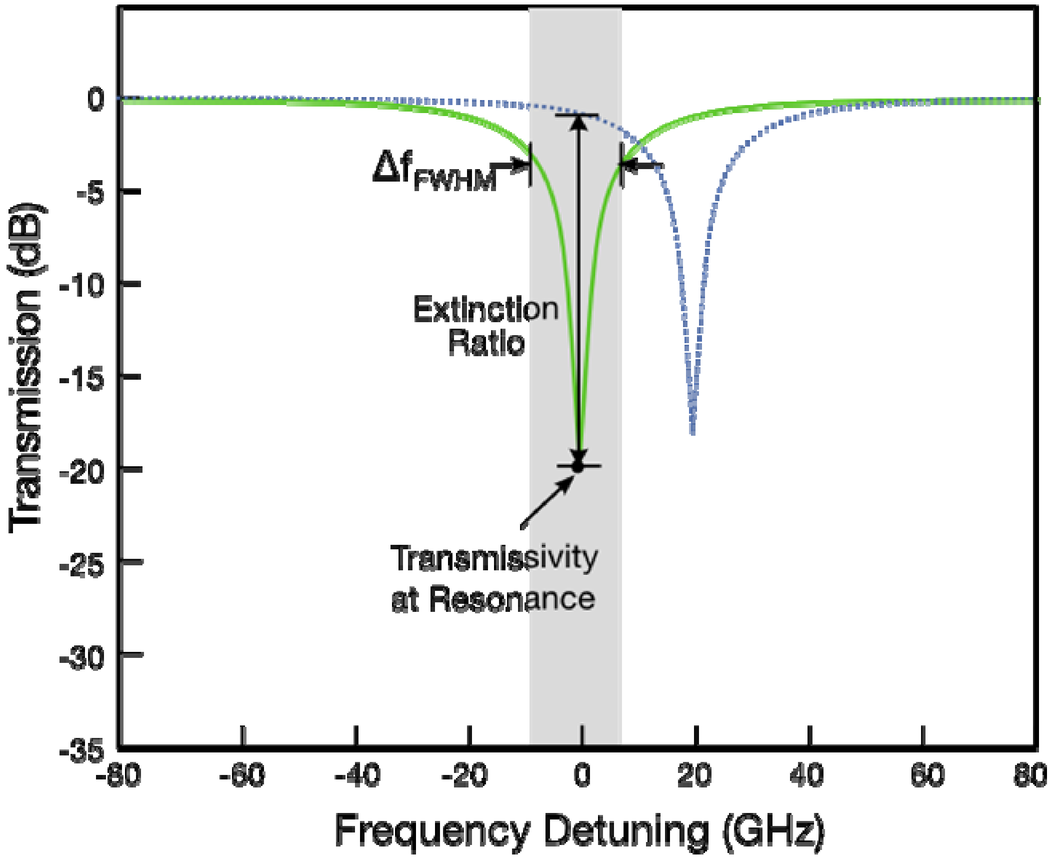
Optical Devices: Ring Modulators



10 Gb/s

Optical Devices: Ring Modulators

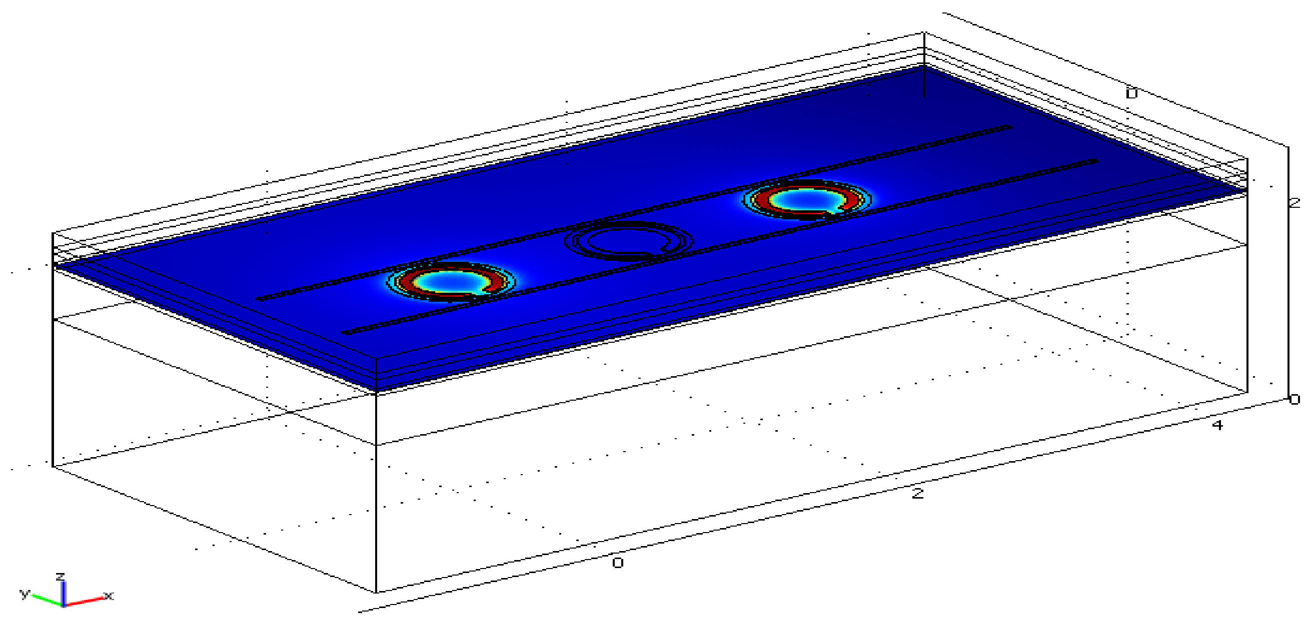




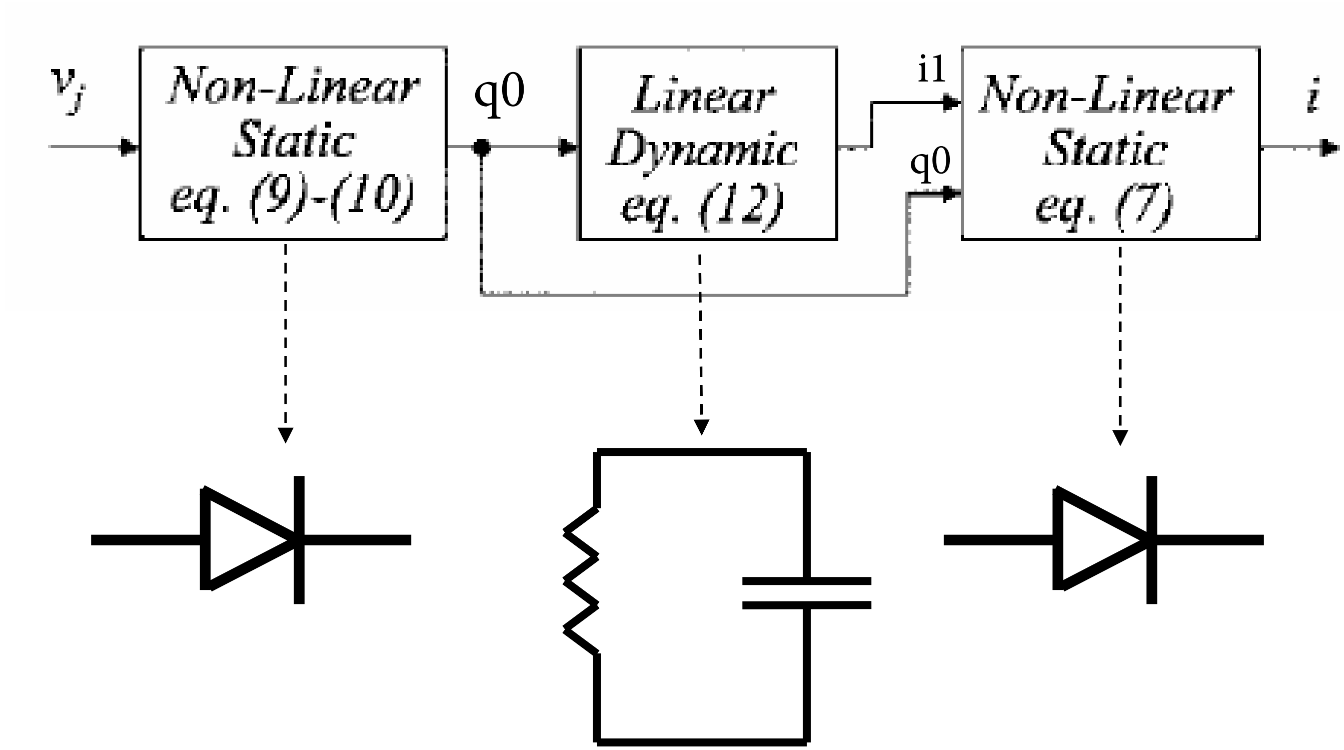
$$T(Q_{injected}) = 1 - \frac{1 - T(0)}{1 + \left(\frac{Q_{injected}}{Q_o}\right)^2}$$

Optical Devices: Ring Modulators

- Thermal crosstalk is high
- Temperature changes ring's refractive index



P-I-N SPICE Model



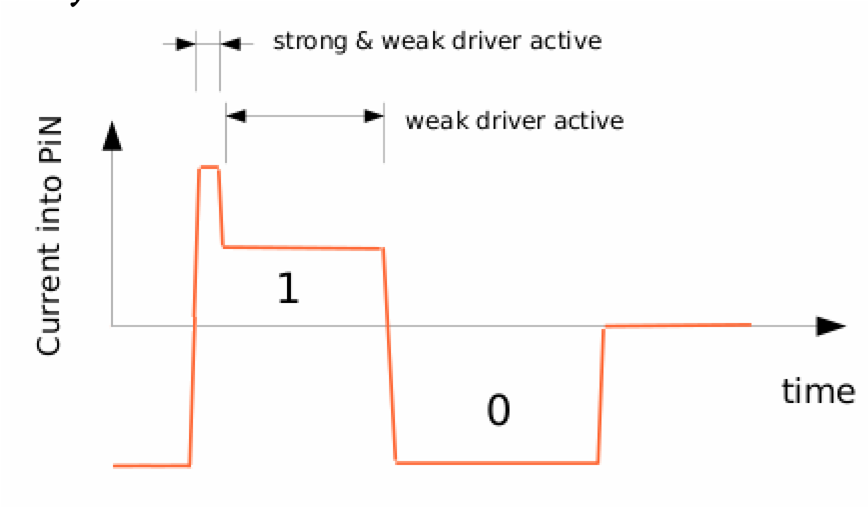
Antonio G. M. Strollo, A New SPICE Model of Power P-I-N Diode Based on Asymptotic Waveform Evaluation, IEEE TRANSACTIONS ON POWER ELECTRONICS, VOL. 12, NO. 1, pp.12-20, JANUARY 1997.

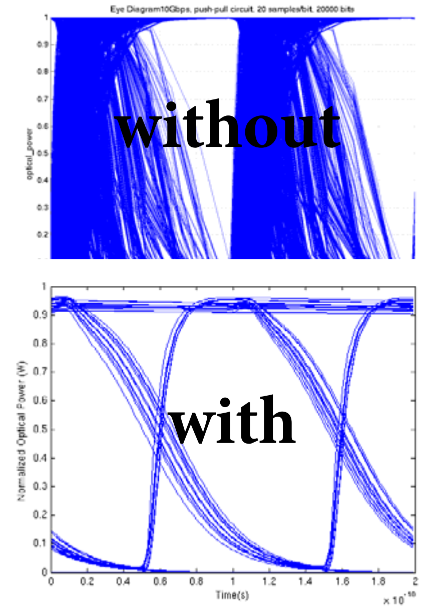
Pre-emphasized Current Profile

Pre-emphasis Advantages

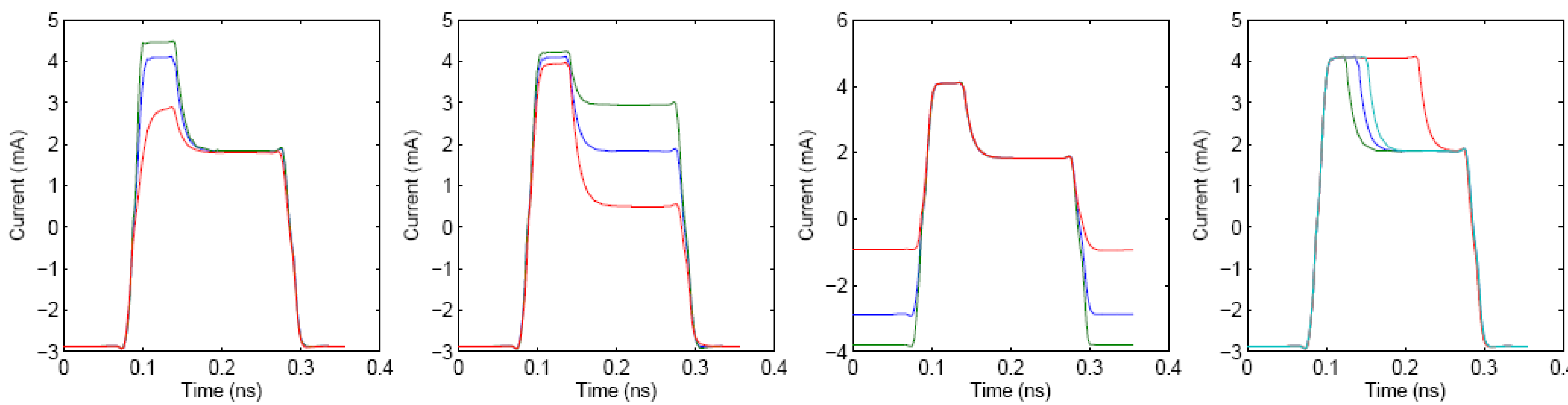
- Less power
- Controlling $Q_{injected}$
- Faster

Injected Current into P-I-N Diode

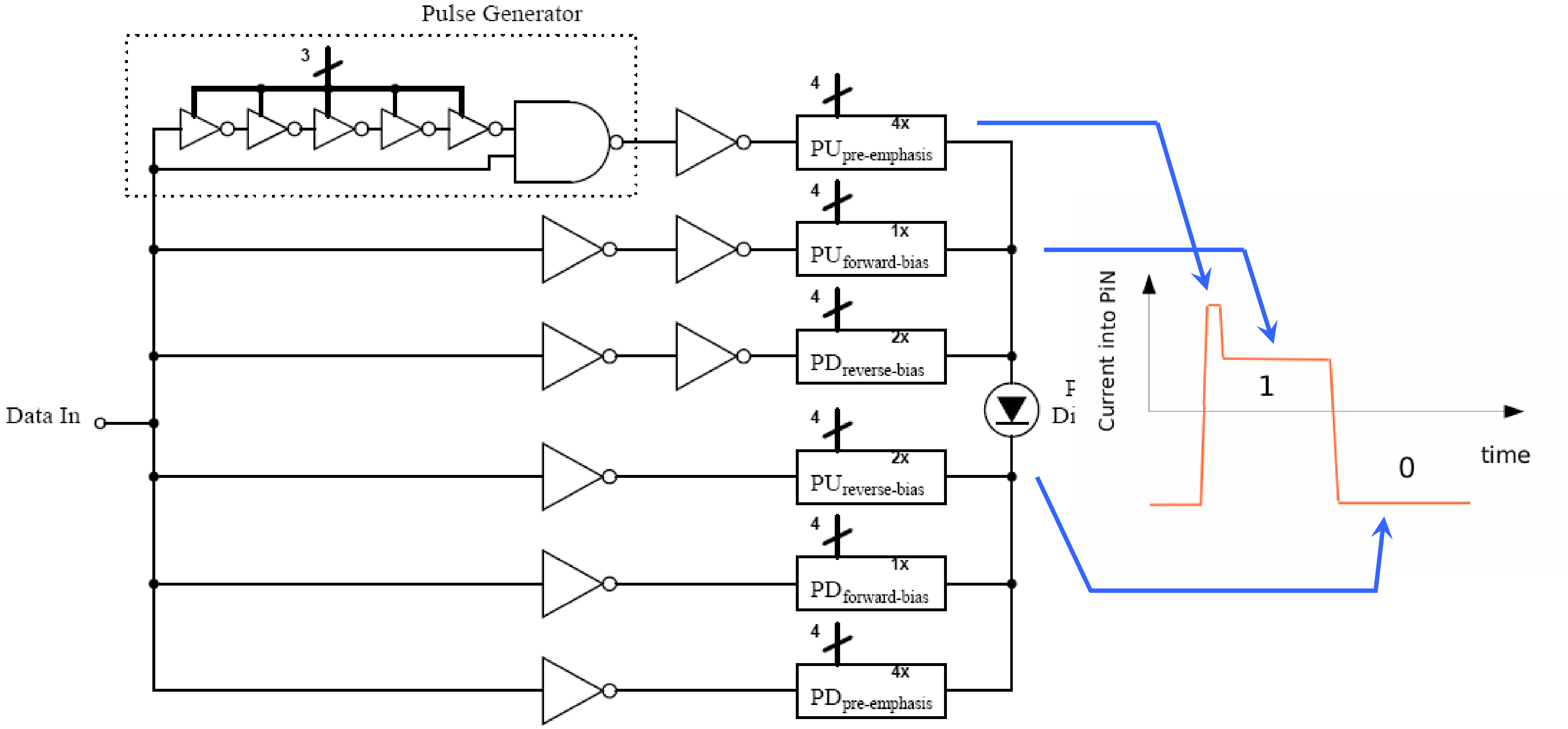




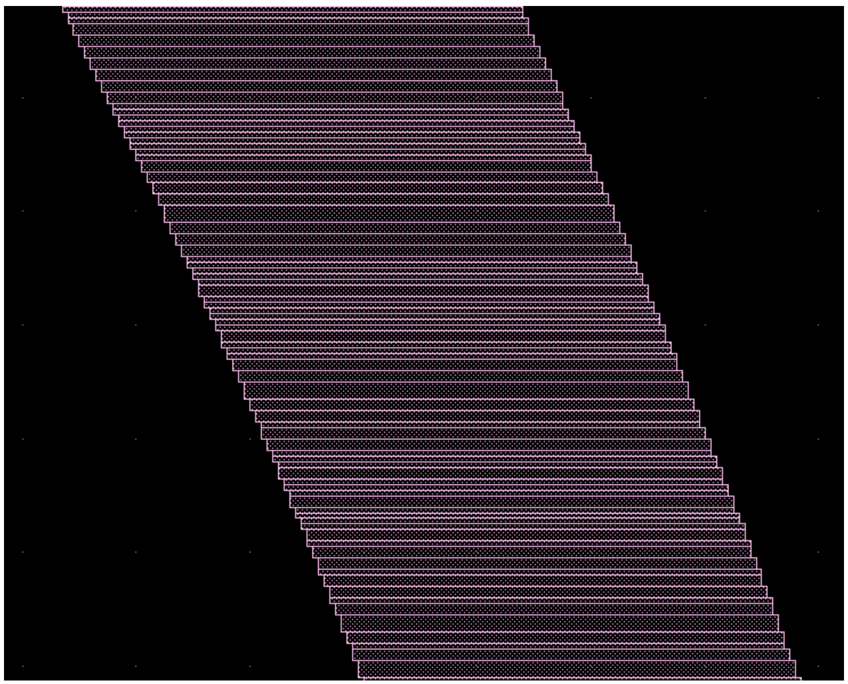
Injected Current Profiles



Modulator Driver Schematic

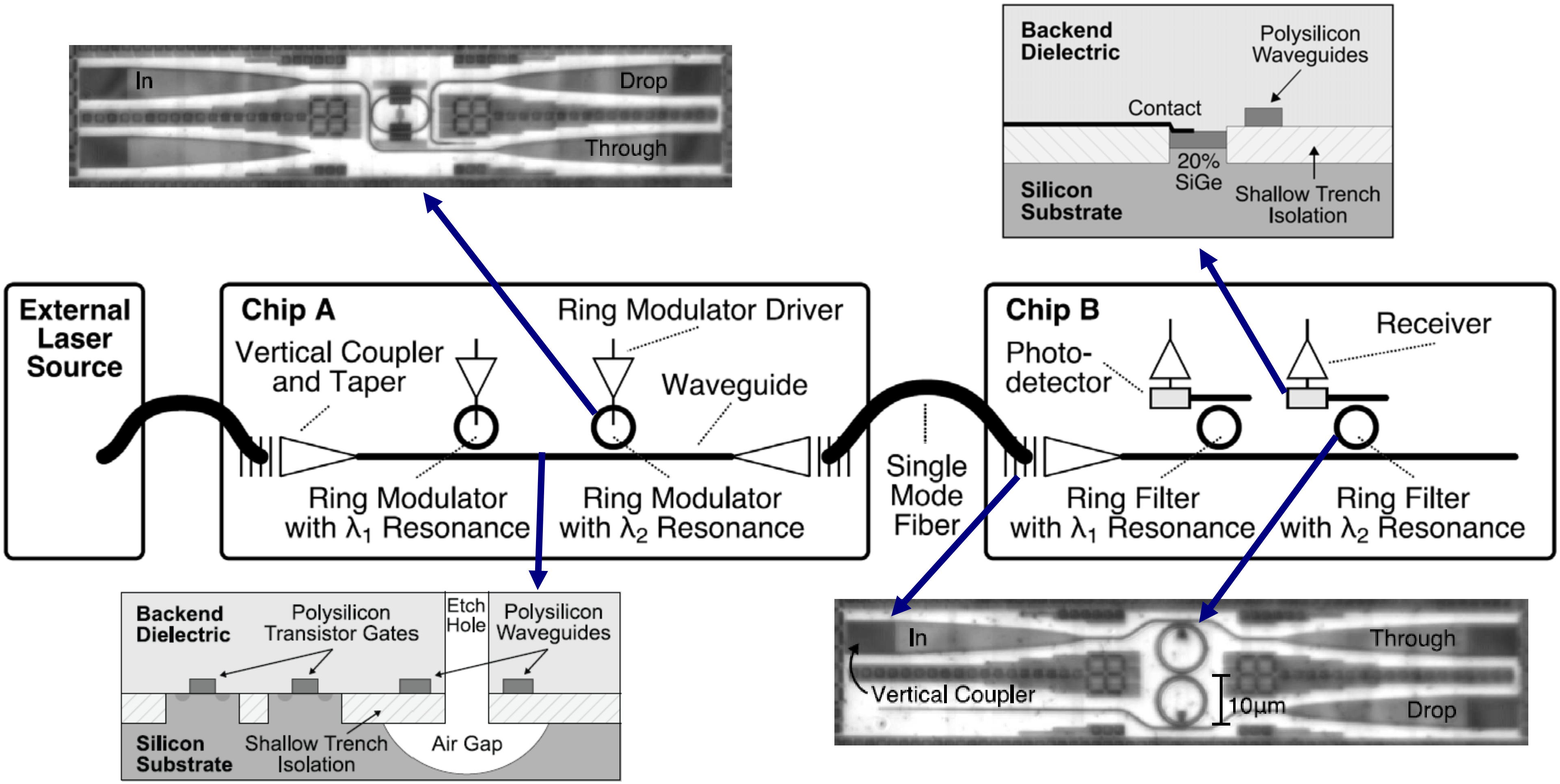


Manufacturing Considerations



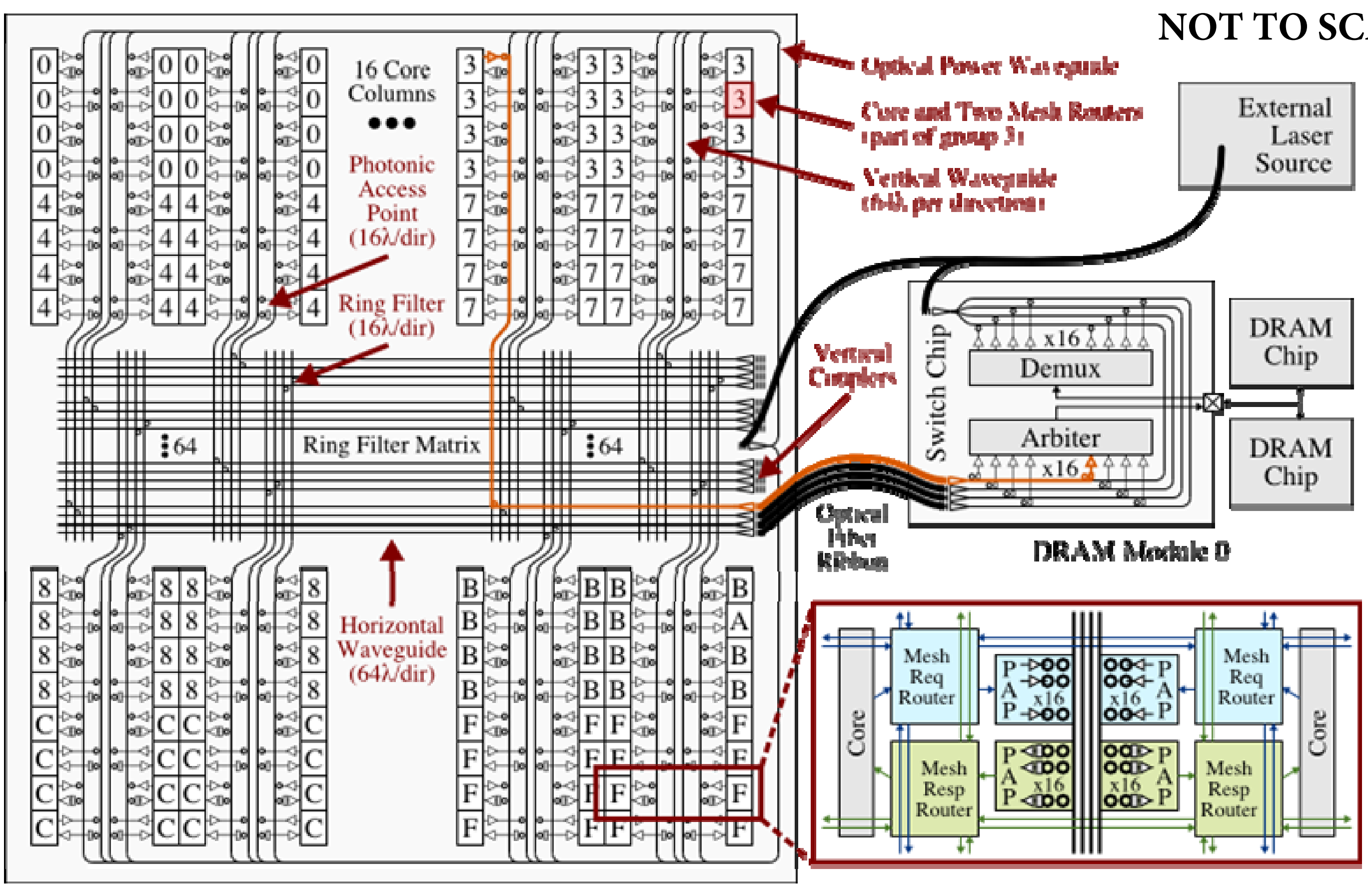
- Manhattan Geometry
- Density requirements
- No Optical LVS
- 1-5 nm design grid
  - Minimizes edge roughness
- Thick-BOX SOI incompatible with processor design (thermal issues)
- Post-processing

An Optical Link



65 nm bulk CMOS chip designed to test various optical devices

Full System (Conventional DRAM)



Photonics < 7% chip area

~40k rings, 200 waveguides, 64λ/waveguide/direction, 40Tbits/sec

Optical power budget

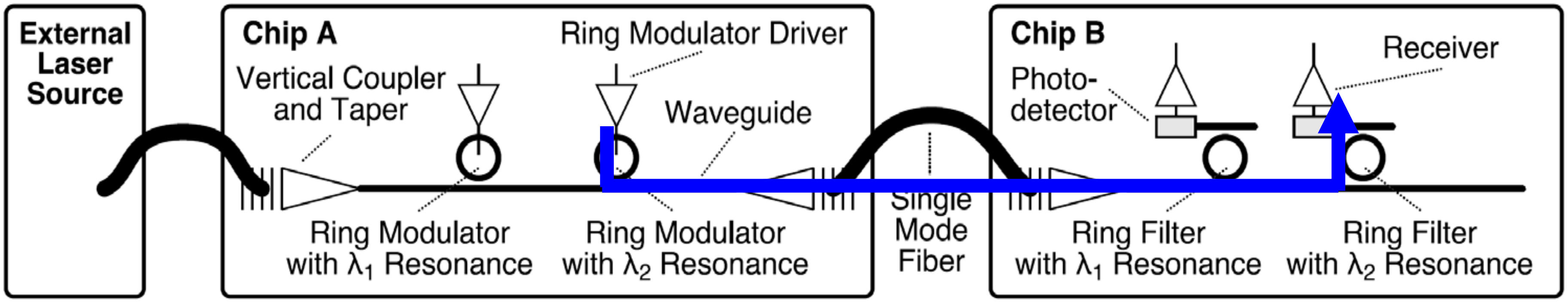
Component	Preliminary Design	Power loss	Optimized Design	Power loss
Coupler loss	1 dB/coupler	3 dB	1 dB/coupler	3 dB
Splitter loss	0.2 dB/split	1 dB	0.2 dB/split	1 dB
Non-linearity	1 dB	1dB	1dB	1dB
Through loss	0.01 dB/ring	3.17 dB	0.01 dB/ring	3.17 dB
Modulator Insertion loss	1 dB	1 dB	0.5 dB	0.5 dB
Crossing loss	0.2 dB/crossing	12.8 dB	0.05 dB/crossing	3.2 dB
On-chip waveguide loss	5 dB/cm	20 dB	1 dB/cm	4 dB
Off-chip waveguide loss	0.5e-5 dB/cm	~ 0 dB	0.5e-5 dB/cm	~ 0 dB
Drop loss	2.5 dB/drop	5 dB	1.5dB/drop	3 dB
Photodetector loss	0.1 dB	0.1 dB	0.1 dB	0.1 dB
Receiver sensitivity	-20 dBm	-20 dBm	-20 dBm	-20 dBm
Power per wavelength		26.07 dBm (0.40 W)		-1.03 dBm (0.78 mW)
Power required at source		**3.3 kW**		**6.38 W**

Data transmission latency

Component	Latency
Serializer/Deserializer (50ps each)	50ps
Modulator driver latency	108ps
Through latency (2.5ps/adjacent channel)	7.5ps
Drop latency (20ps/drop)	60ps
Waveguide latency (106.7ps/cm)	427ps
SM fiber latency (48.3ps/cm)	483ps
Photodetector+TIA latency	200ps
Total latency	1.385ns

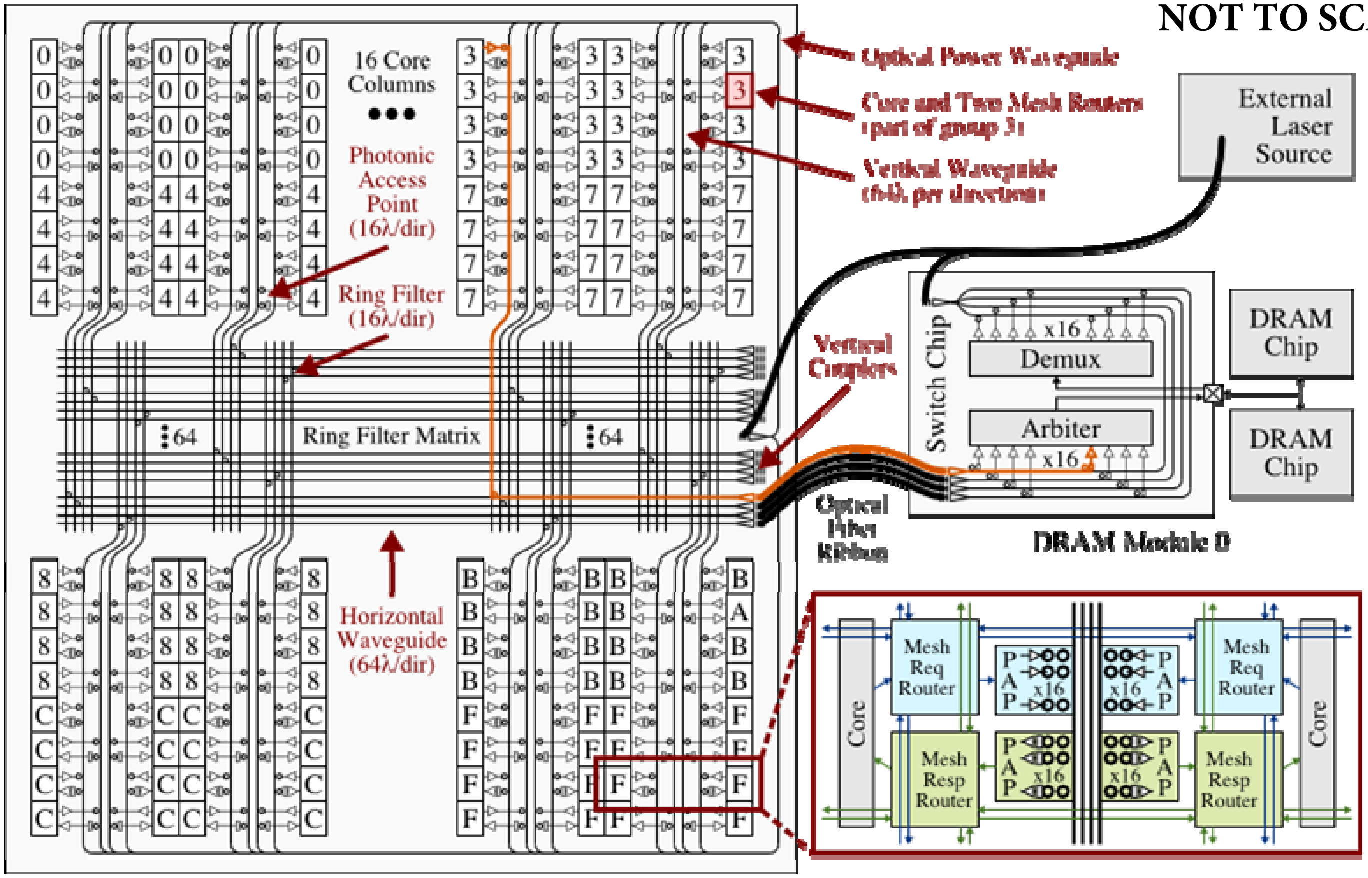
- Total latency 14 bit times
  - Less than 4 clock cycles (with 2.5 GHz core clock)
- Comparable to Electrical

Silicon Photonics Area and Energy Advantage



Metric	Energy (pJ/b)	Bandwidth density (Gb/s/μm)
Global on-chip photonic link	0.25	160-320
Global on-chip optimally repeated electrical link	1	5
Off-chip photonic link (50 μm coupler pitch)	0.25	13-26
Off-chip electrical SERDES (100 μm pitch)	5	0.1
On-chip/off-chip seamless photonic link	0.25	

Thermal Sensitivity/Crosstalk Issues



- High thermal crosstalk between adjacent rings
- Released substrate can thermally isolate filter banks
- Random variation small; systematic variation corrected thermally

The Problem Spans Many Layers

- **Device Level** – Jason Orcutt, Milos Popovic, Anatoly Khilo, Charles Holzwarth, Jie Sun, Hanquing Li, Reja Amatya
- **Circuits Level** – Ben Moss, Michael Georgas, Jonathan Leu
- **System Architecture Level** – Ajay Joshi, Imran Shamim, Chris Batten

Conclusion

- Multi-core machines demand more bandwidth than ever before
- On-chip optical networks are a new technology to widen this bottleneck
- We have demonstrated that on-chip optical networks are possible in standard CMOS processes
- Many tradeoffs exist that must be analyzed at the system, circuit and device level
- Photonic interconnects have over 5x energy efficiency improvement compared to electrical