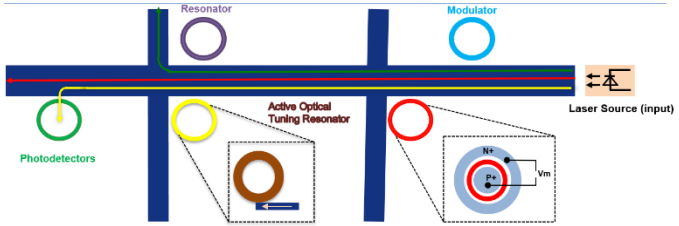
Si-Photonics Technology Towards fJoule/bit Optical Communication in Many-core Chips

Adaptive Systems Lab
Abderazek Ben Abdallah
benab@u-aizu.ac.jp

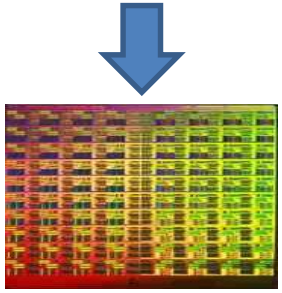


Contents

Supercomputer (1996) Vs. Intel Tera-scale CPU (2007)

ASCI Blue-Mountain (1.6 TeraOps, 929 m^2, 1.6 Mwatts)

Ref. http://www.jipdec.or.jp



Intel's Tera-scale 80 core Chip

(1.63 Teraflops @ 5.1 GHz, 175 watts, and 1.81 Teraflops @ 5.7 GHz, 265 watts).

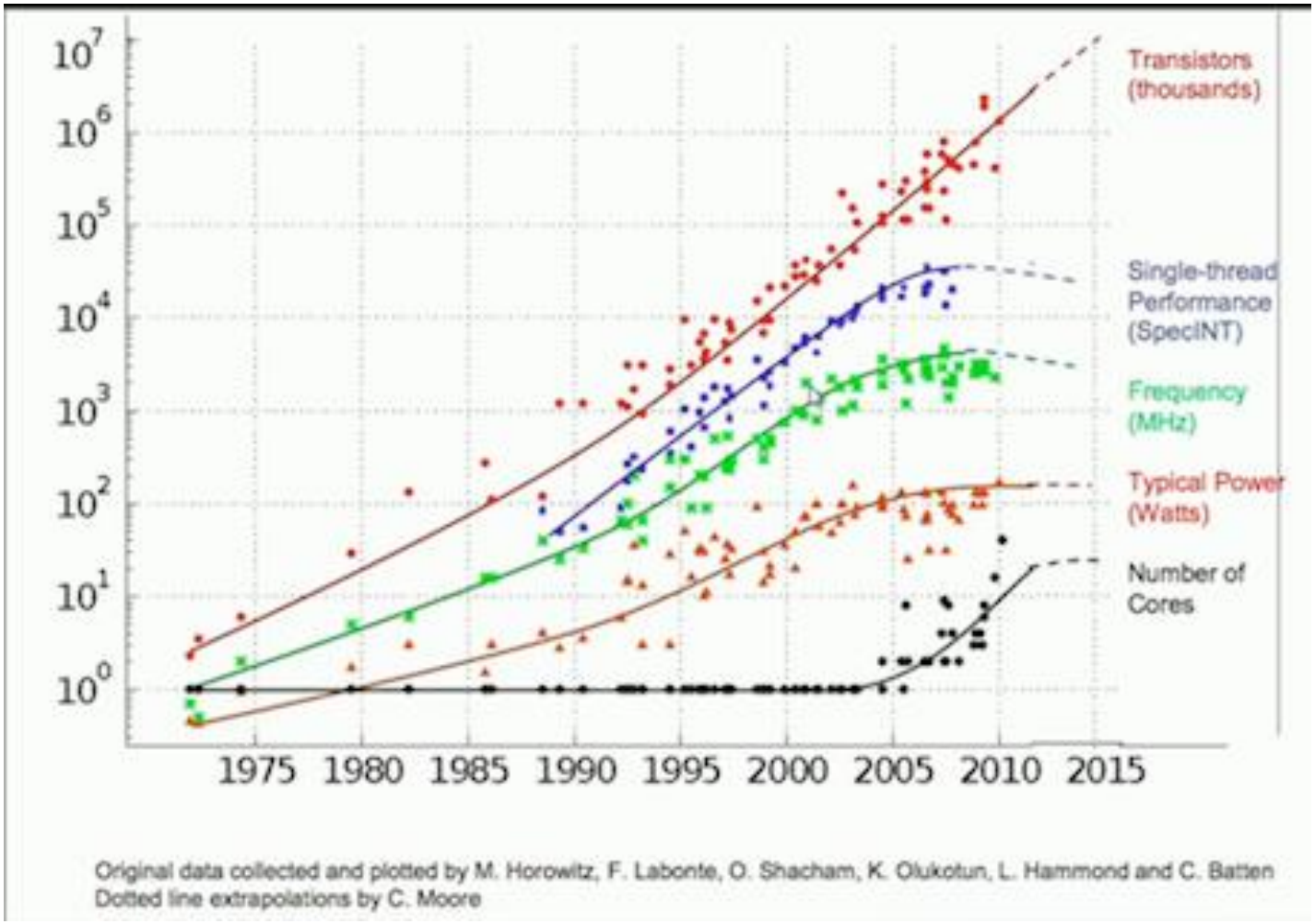
Supercomputer (1996) Vs. Sony PS (2006)

ASCI Blue-Mountain: 1.6 TeraOps, 929 $m^{2,}$ 1.6 Mwatts

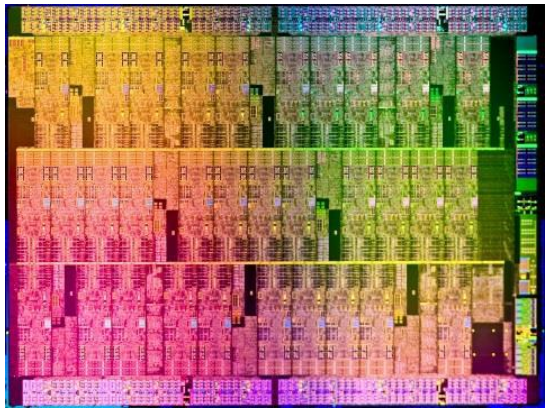
Ref. http://www.jipdec.or.jp

Sonny Play station 3 (2006): 1.8 Teraflops, 0.08 square meter, <200 watts

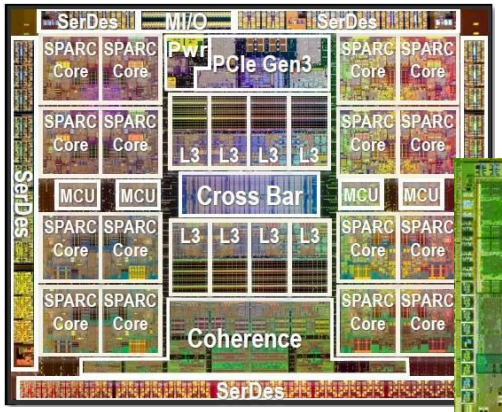
Moore's Law


10^7
10^6
10^5
10^4
10^3
10^2
10^1
10^0
1975 1980 1985 1990 1995 2000 2005 2010 2015
Transistors (thousands)
Single-thread Performance (SpecINT)
Frequency (MHz)
Typical Power (Watts)
Number of Cores
Original data collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond and C. Batten
Dotted line extrapolations by C. Moore

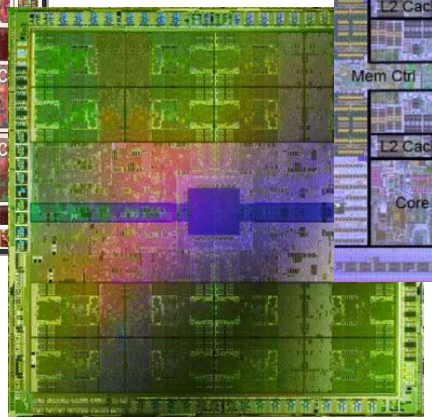
Current Processor Research Trends



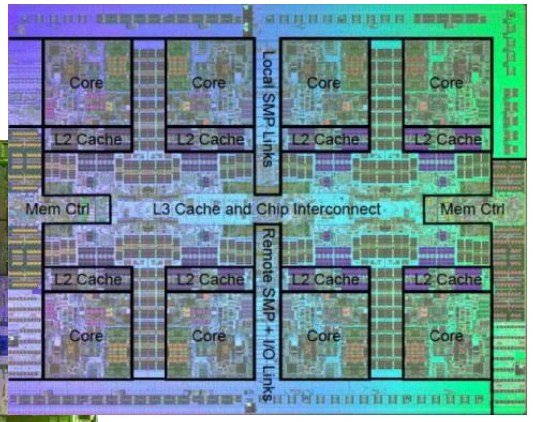
Intel Knights Corner 50 cores, 200 Threads



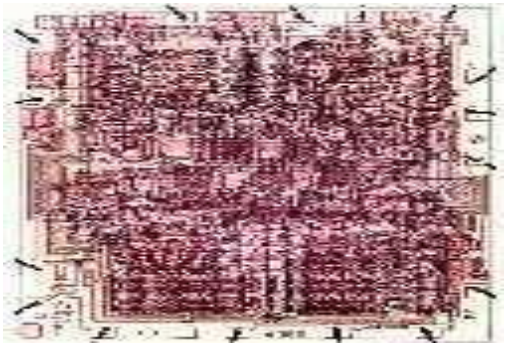
Oracle T5 16 cores, 128 Threads



Nvidia Fermi 540 CUDA cores



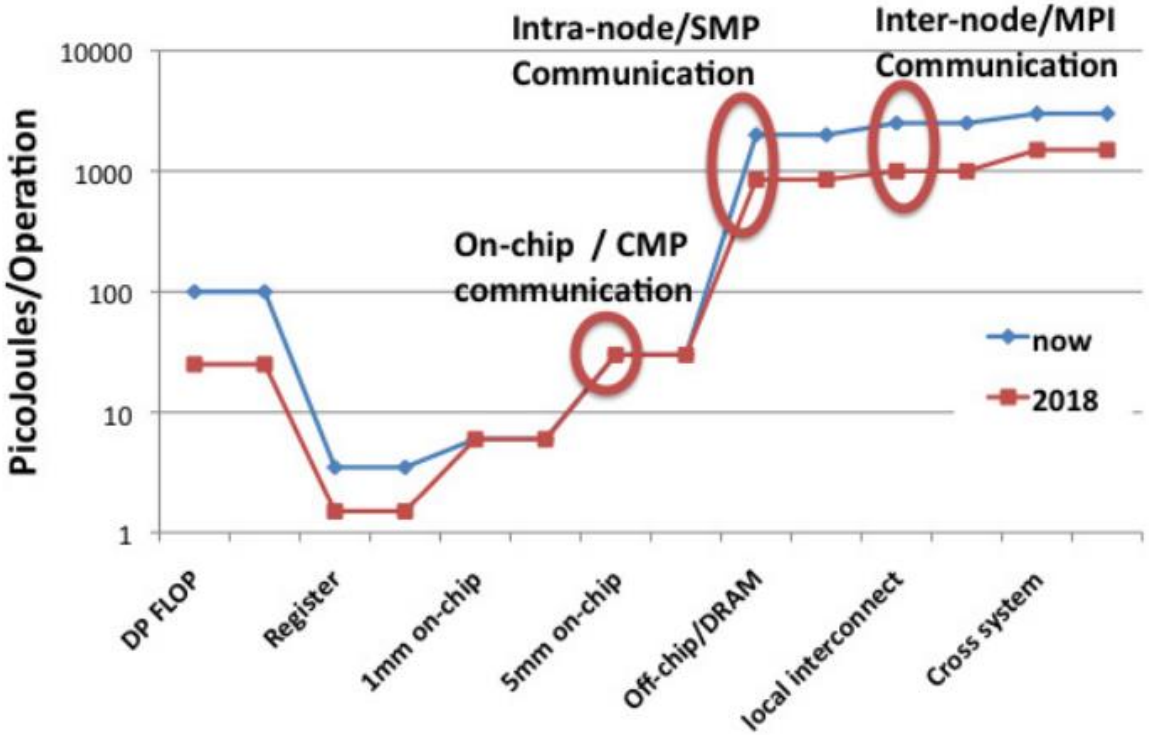
IBM Power 7
8 cores, 32 threads



Intel 4004 (1971): 4-bit processor, 2312 transistors, ~100 KIPS, 11 mm2 chip

1000s of processor cores per die could be integrated?
→ How about power scaling!

Wire and I/O scaling problems

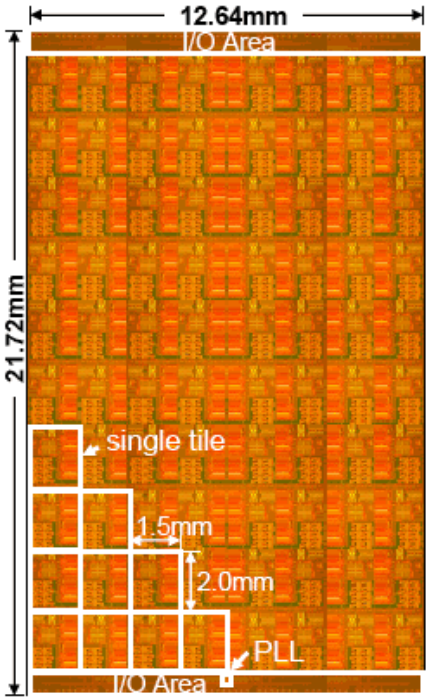


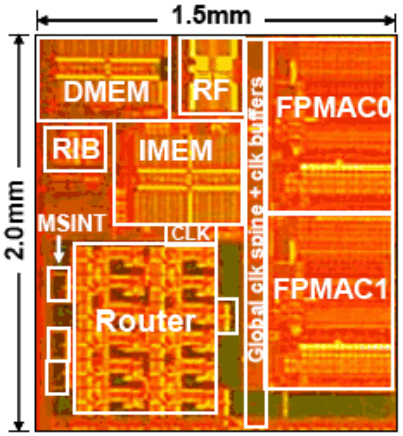
Energy cost of data movement relative to the cost of a flop for current and 2018 systems. (Shalf et al., VECPAR 2010)

- Preparing the operands costs more than performing computing on them!
- There is no Moore's law for communications!

Current Processor Research Trends

-Easier Programming
-Easier Implementation
-Low energy efficiency
-No specific HW for different tasks.



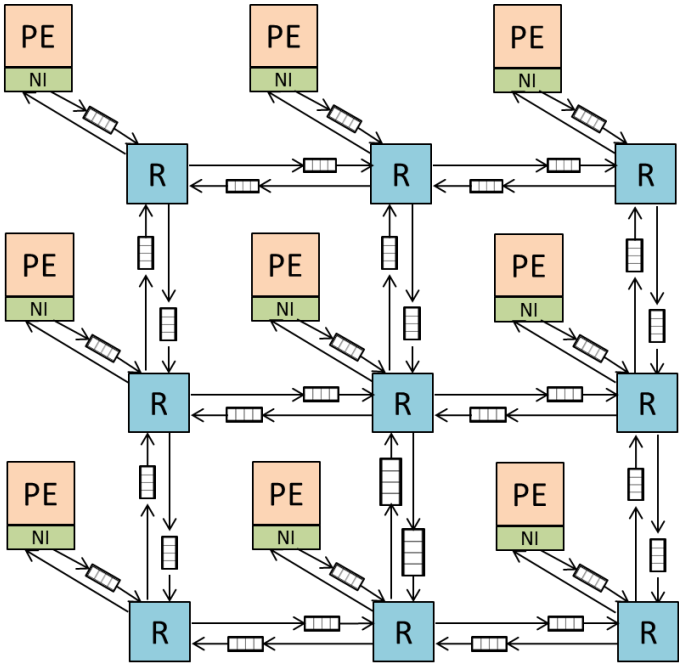


Technology	65nm CMOS Process
Interconnect	1 poly, 8 metal (Cu)
Transistors	100 Million
Die Area	275mm²
Tile area	3mm²
Package	1248 pin LGA, 14 layers, 343 signal pins

Intel 80-core (Homogeneous System)

- Conventional Electric-wiring on chip (add-hoc wiring) consumes half of CPU power.
- Teraflop Chip router consumes 28% of CPU power.

Limitations of Traditional E-NoC

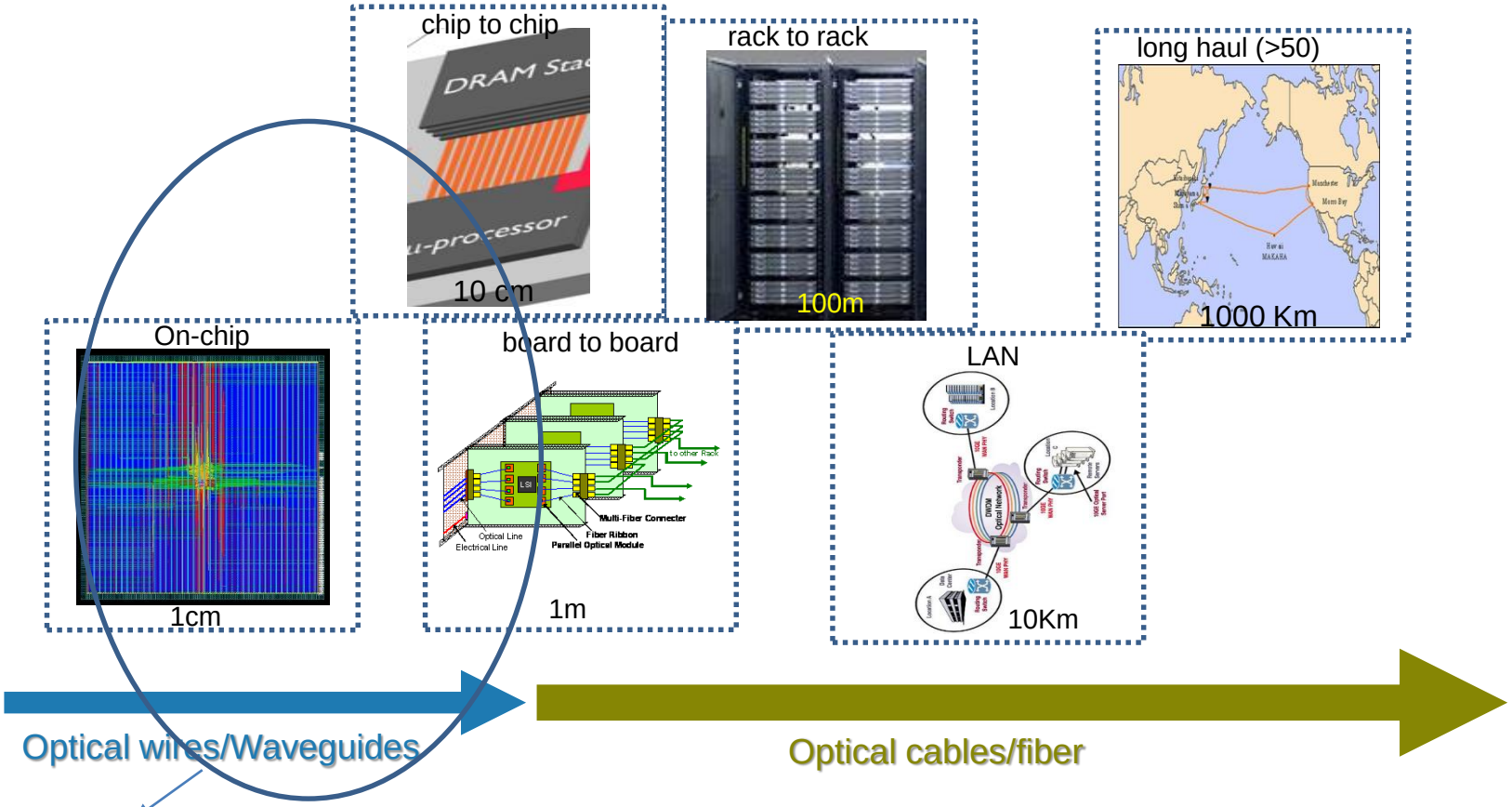


R: Router
NI: Network interface
PE: Processing Element

- Multi-hop communication.
- Receive, buffer and retransmit every bit at every switch.
- High latency and energy dissipation especially in large system.

Limitation of Electric/Metal wire

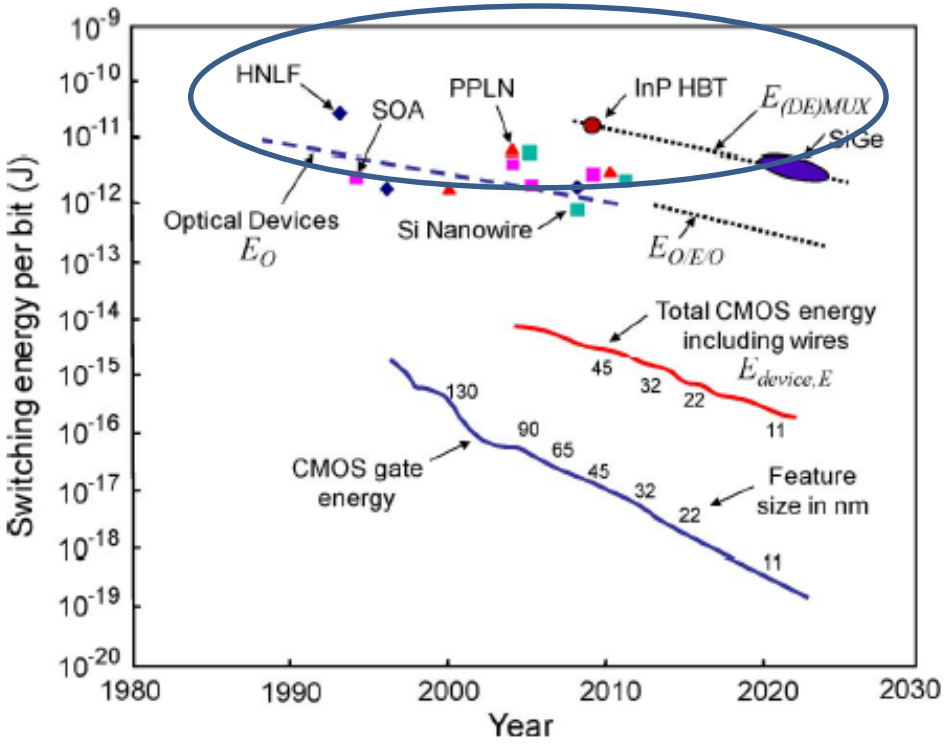
- Electronics is not good at high bit/s communication.



Can we have photonic networking?

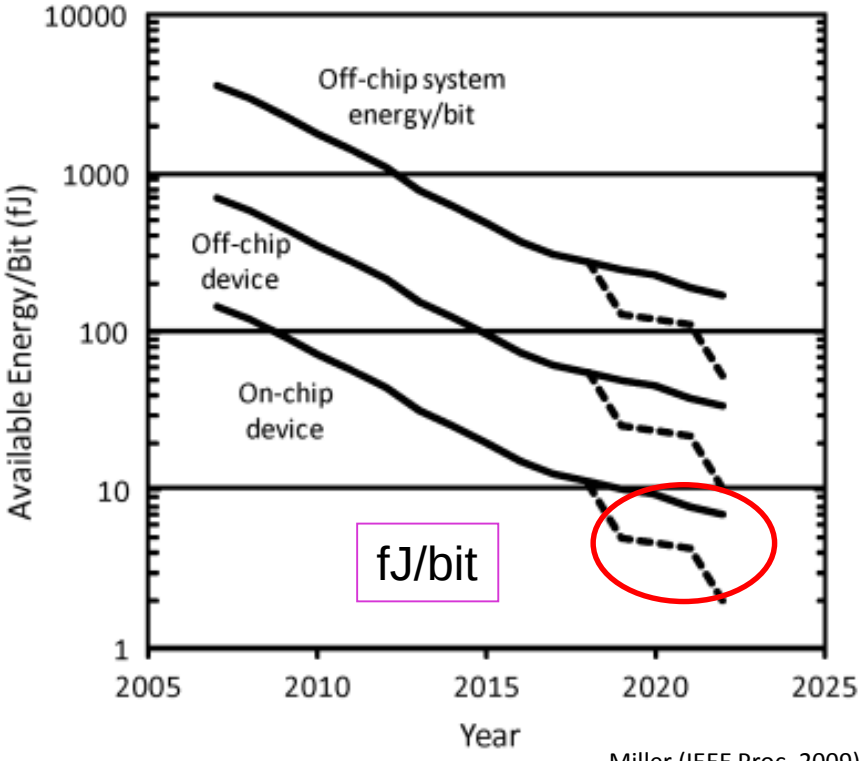
Energy Cost for Communication

Conventional photonics



Tucker (IEEE Photonics Jnl 2011).

Required energy cost



Miller (IEEE Proc. 2009)

Small transmission energy, but high processing energy.

We need a "Spring-Revolution"to deal with the Power/Energy Wall!

- The computation power of CPU is still progressing exponentially, and there are strong demands to keep this progress rate for the next decades.

- If we assume the same progress rate, the allowable energy for transmitting a single bit in a chip should be around a few fJ in 2025 [Miller 2009].

- The problem of on-chip electric communication is largely attributed to the finite RC of wirings.
  - As the bit rate goes up, we have to use wider and shorter wires in order to avoid the RC delay → Conflicting with the limited space-budget in a chip!!.

Contents

- 1. Trends in CPU
- **2. Optical interconnect**
- 3. Si-Photonics Many-core chips
- 4. Research direction, challenges
- 5. Concluding remarks

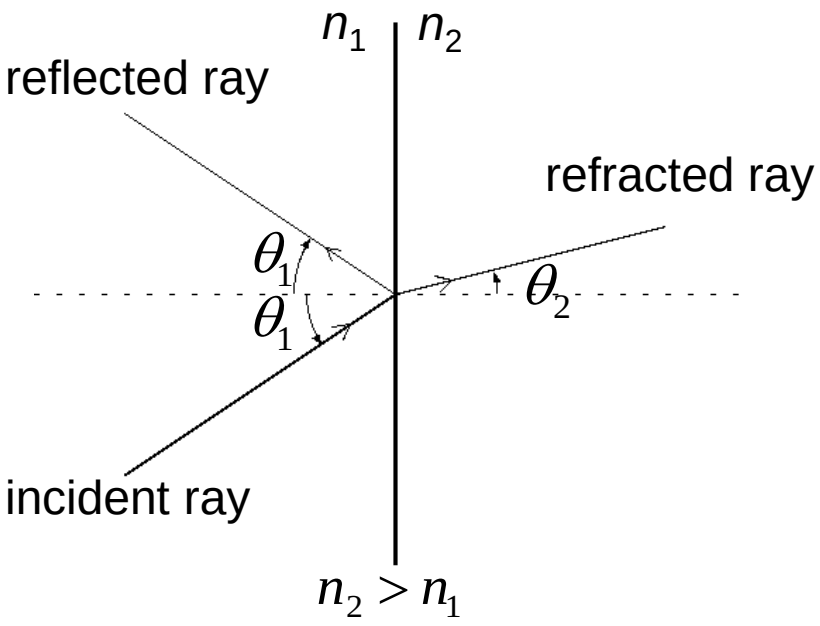
Why optical interconnection ?

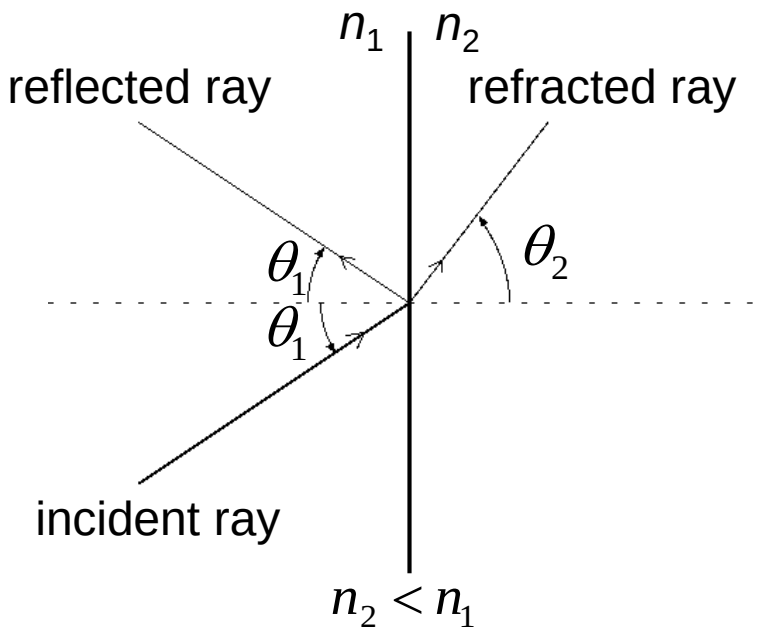
- **Larger bandwidth is possible for a long wire**
  - Bandwidth can be enhanced by WDM.
- **Efficient Energy at high bit rate communications.**
  - No energy cost for transfer (no charging energy)
- A photon can generate ≈1 volt (via photo-electric effect), which is NOT bound by the light intensity (number of photons).

Transmission over Si Wire/Waveguide

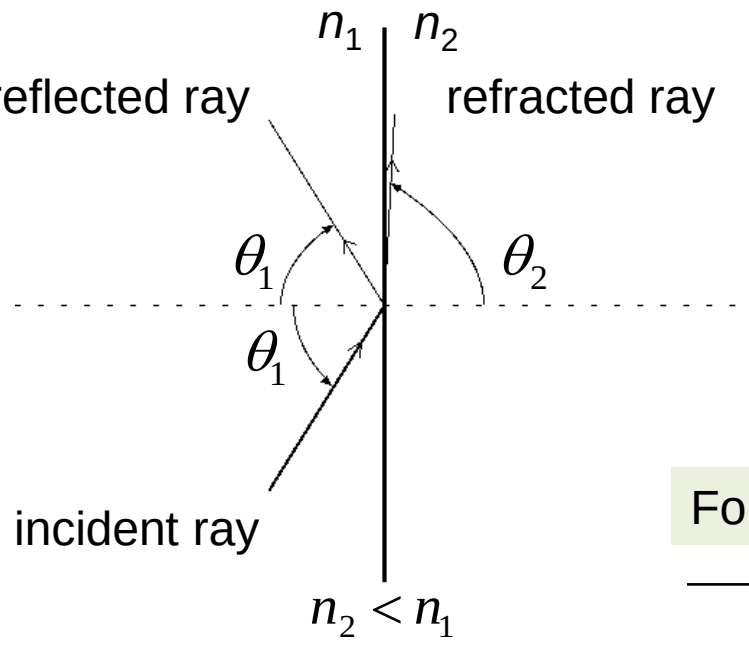
Snell's Law of Refraction:

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{n_2}{n_1} = \frac{v_1}{v_2}$$





Total internal reflection in Si Wire/Waveguide



Let $\theta_2 = \pi/2$:

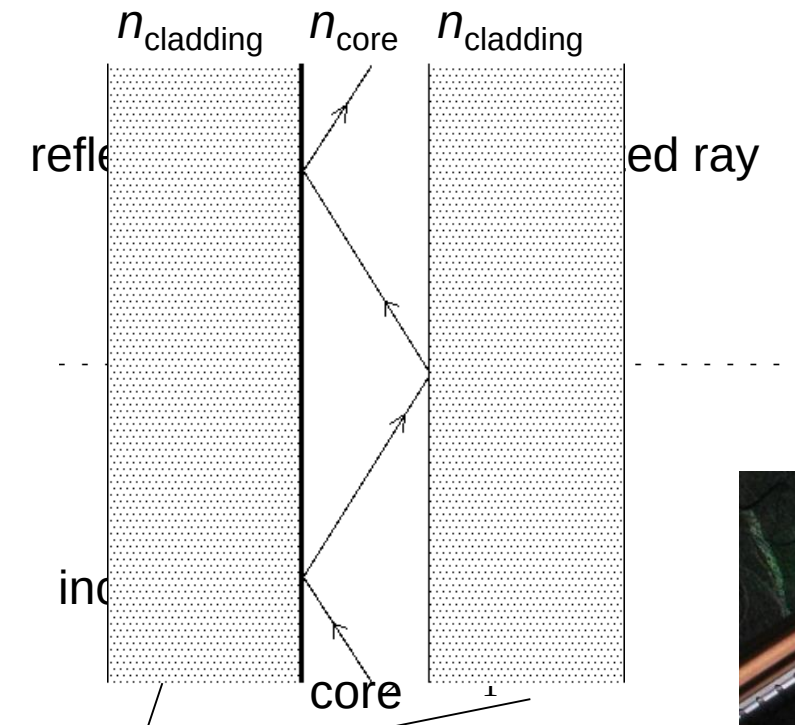
Then $\sin\theta_1 = \dfrac{n_2}{n_1}$

$$\longrightarrow \theta_c = \sin^{-1}\left(\frac{n_2}{n_1}\right)$$

For $\theta_1 > \theta_c$, light ray is completely reflected.

$\longrightarrow$ *Total internal reflection*

Total internal reflection in Si Wire/Waveguide



$n_{cladding} < n_{core}$

Total internal reflection keeps all optical energy within the core, even if the fiber bends.

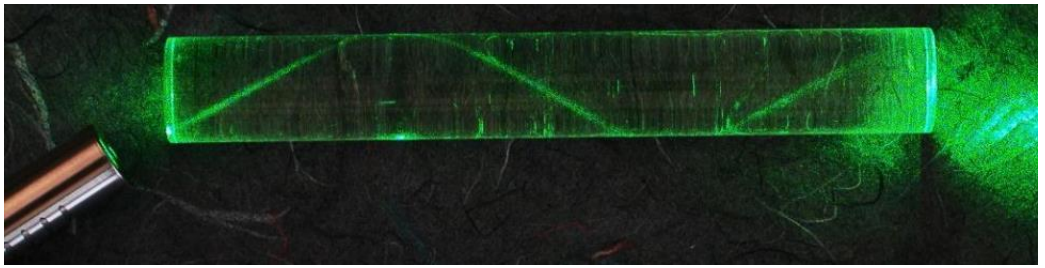
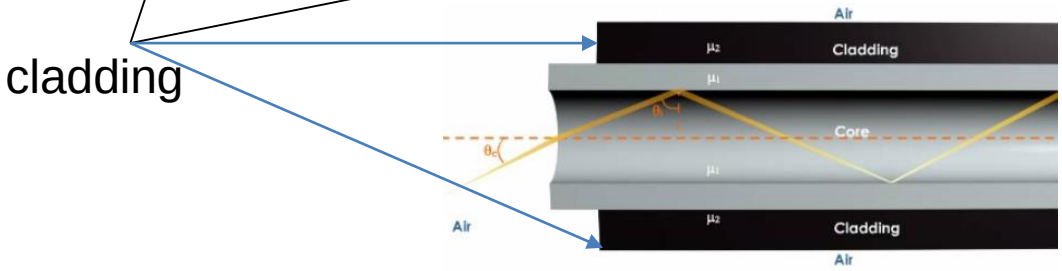
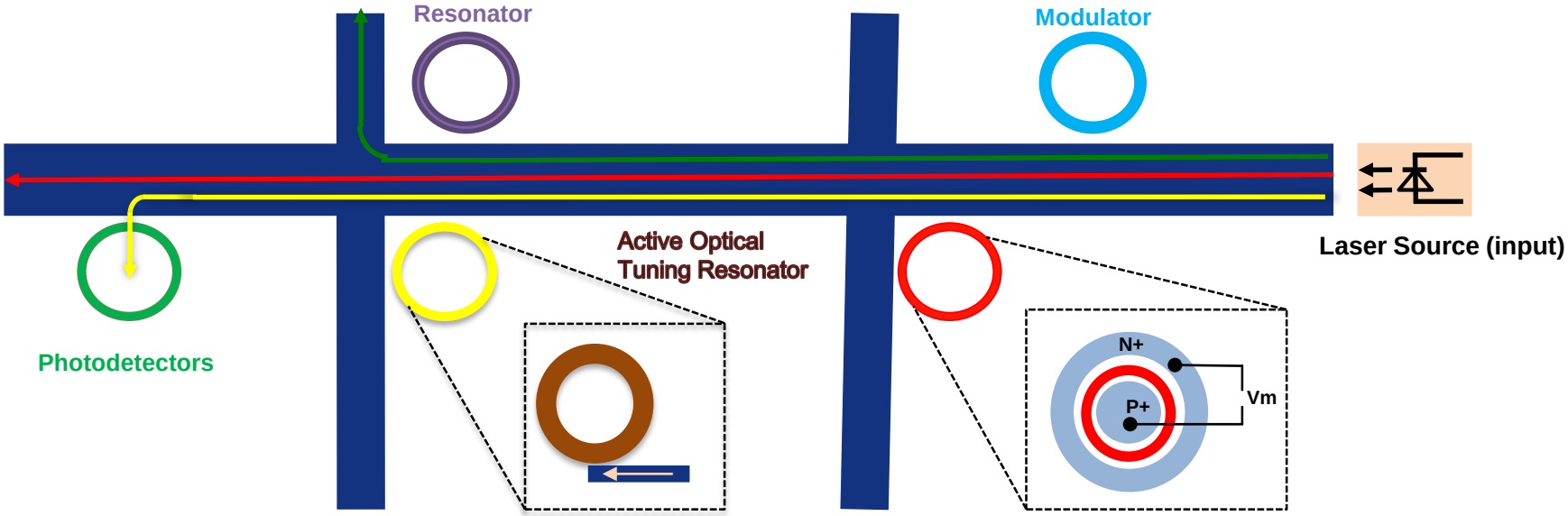


image from Wikipedia



Si-Photonics building blocks



Main components

- **Laser Source**: Inject the required laser lights into waveguide
- **Modulators**: Modulate the laser lights to ‘0’ and ‘1’ states
- **Photodetectors**: Detect the laser lights and convert to electrical signal
- **Turn Resonators**: Control the routing direction of the laser lights

Problems in Photonic Integration

- Fabrication cost → *Being explored by Si photonics.*
- Low energy cost for data transmission → This is *a big issue.* *How much should we reduce ?*
- Larger scale with higher density → What applications for *large-scale photonics ?*

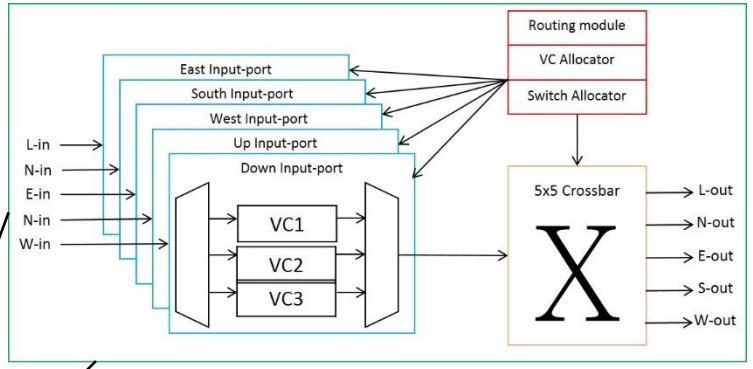
Contents

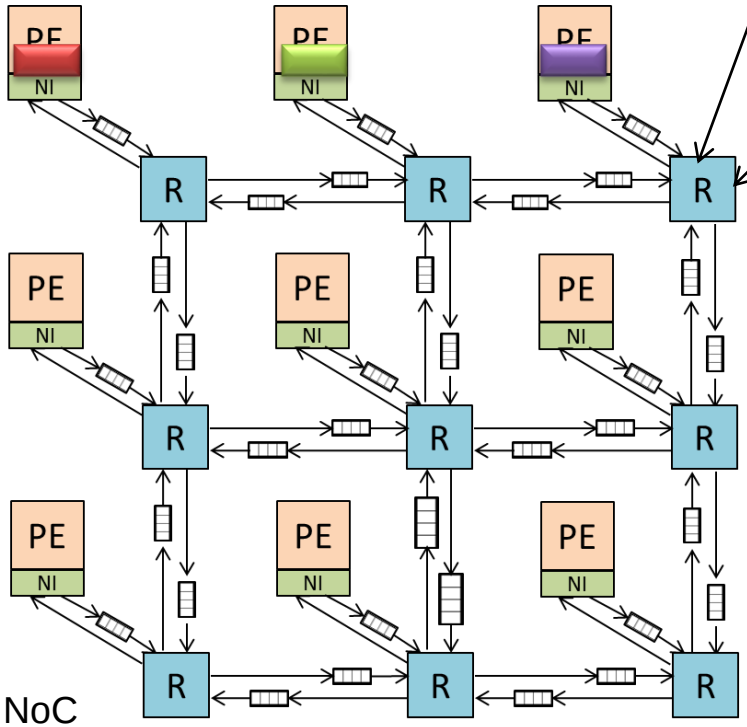
- 1. Trends in CPU
- 2. Optical interconnect
- **3. Si-Photonics Many-core chips**
- 4. Research direction, challenges
- 5. Concluding remarks

OASIS-1:
Overview of Electronic Packet Switched NoC

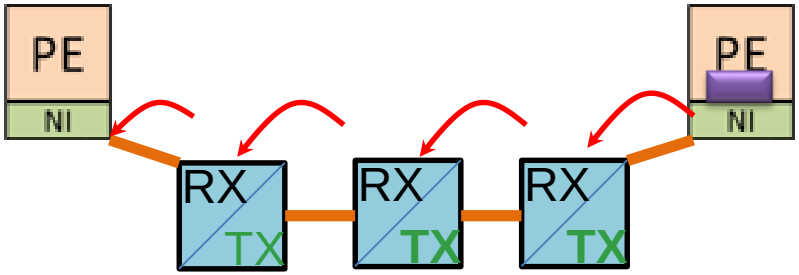
Typical Packet format

Tail flit	Body flit	Head flit
Ending flit	Carried Payload	Flit information





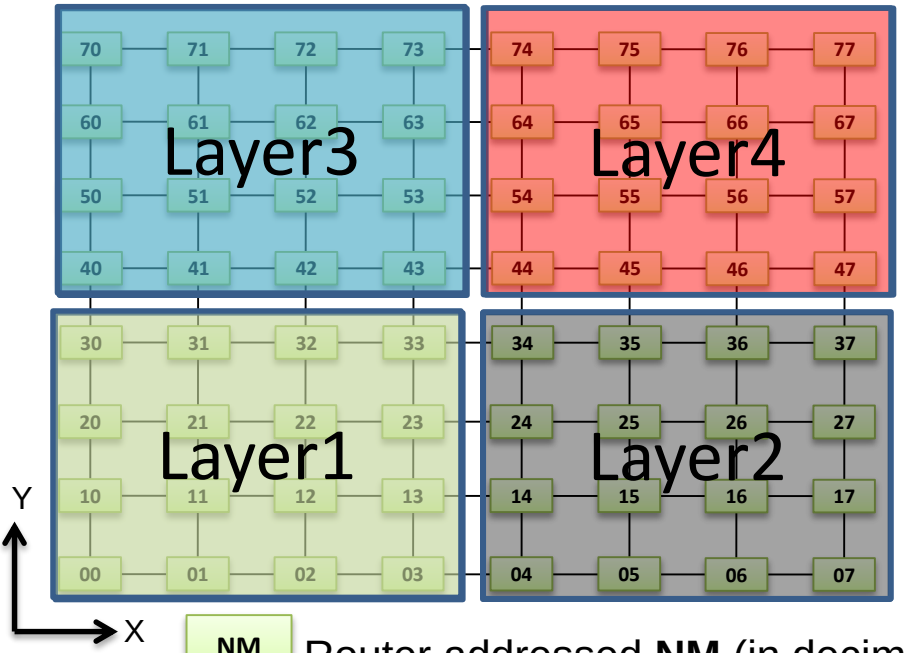
Scalability issue if chip is very large → Latency, bandwidth, and power problems.



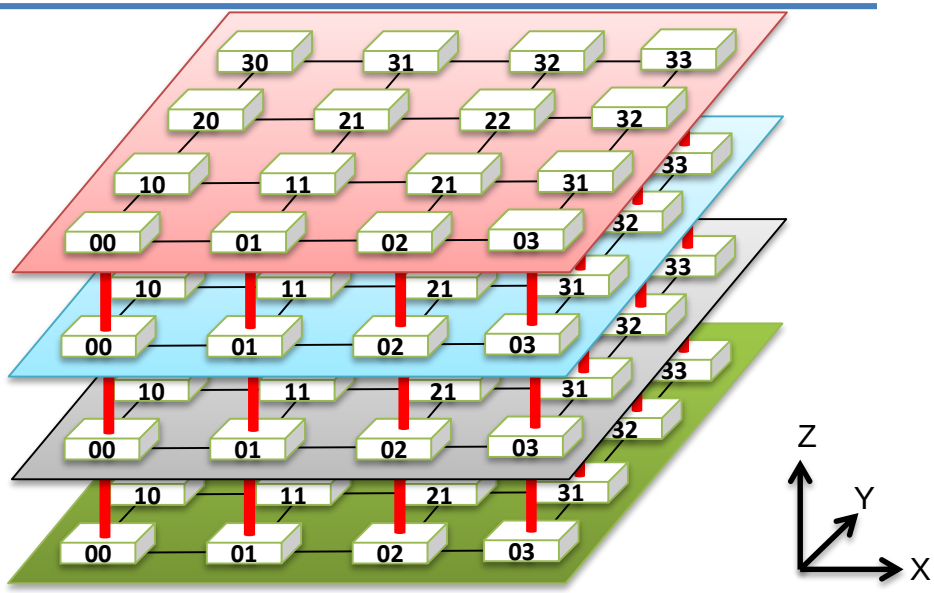
Multihop communication
Receive -> Buffer - > Transmit
every flit at every switch.

R: Router. NI: Network interface. PE: Processing Element

OASIS-1:
Overview of Electronic Packet Switched NoC

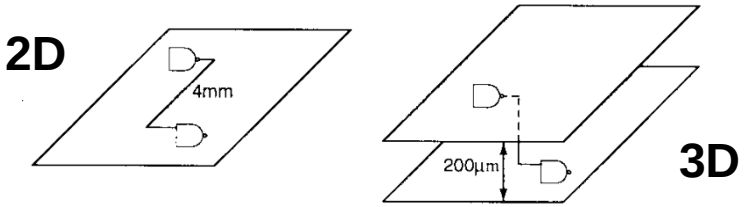


NM Router addressed **NM** (in decimal)

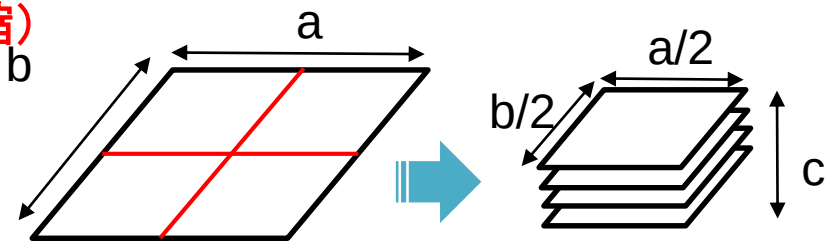


3D- Network-on-Chip architecture

Wire length reduction（配線短縮）

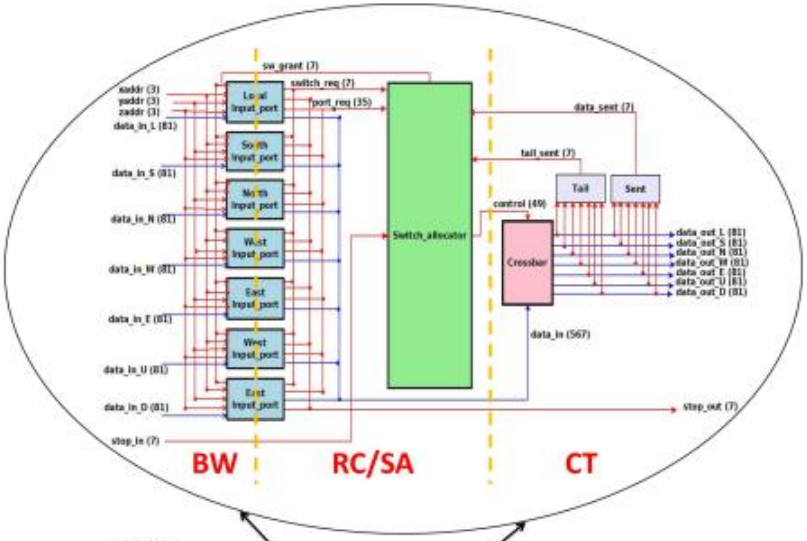


Footprint reduction（面積削減）



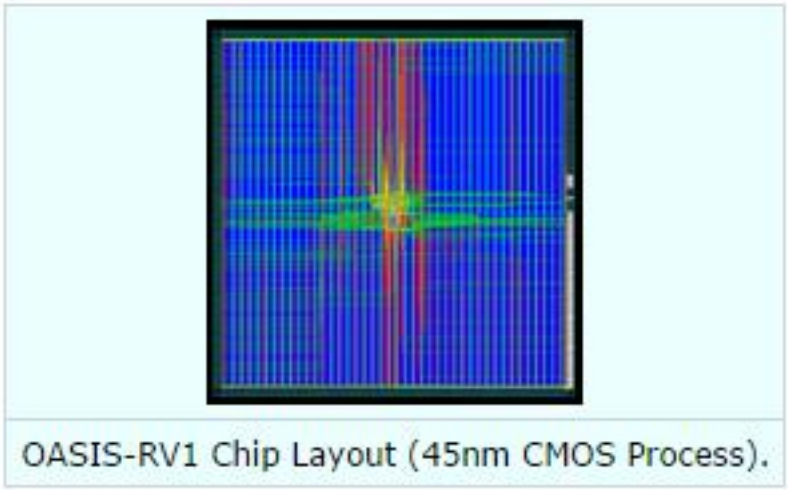
Lateral link (1mm ~ 4mm)
Vertical link (10 µm ~200 µm)

OASIS-1:
Overview of Electronic Packet Switched NoC



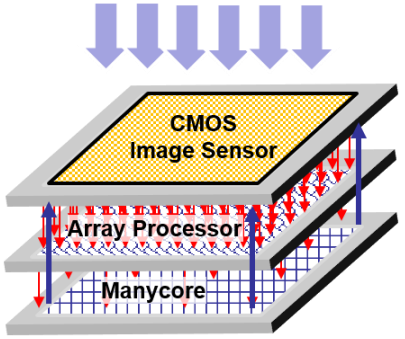
OASIS Network-on-Chip System

(7/14)



OASIS-RV1 Chip Layout (45nm CMOS Process).

Power: 222.387 uW, Number of Pins: 557

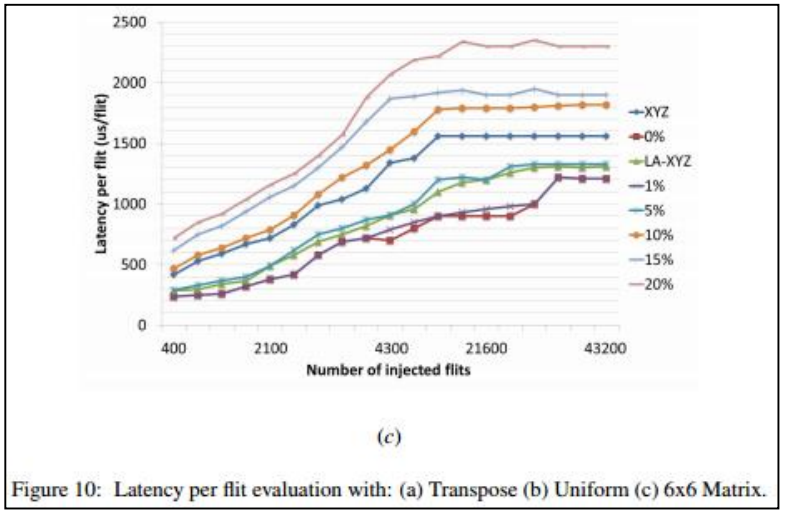


Credit: Y. Matsumoto

OASIS-1:
Overview of Electronic Packet Switched NoC

Table 1: Simulation configuration.

Parameters / System		LAFT	LA-XYZ	XYZ
Network Size	Matrix	3x6x6	3x6x6	3x6x6
Mesh	Transpose & Uniform	4x4x4	4x4x4	4x4x4
flit size		34 bits	34 bits	31 bits
Header size		13 bits	13 bits	10 bits
Payload size		21 bits	21 bits	21 bits
Buffer Depth		4	4	4
Switching		Wormhole-like	Wormhole-like	Wormhole-like
Flow control		Stall-Go	Stall-Go	Stall-Go
Scheduling		Matrix-Arbiter	Matrix-Arbiter	Matrix-Arbiter
Routing		Look Ahead Fault Tolerant	Look-Ahead-XYZ	XYZ

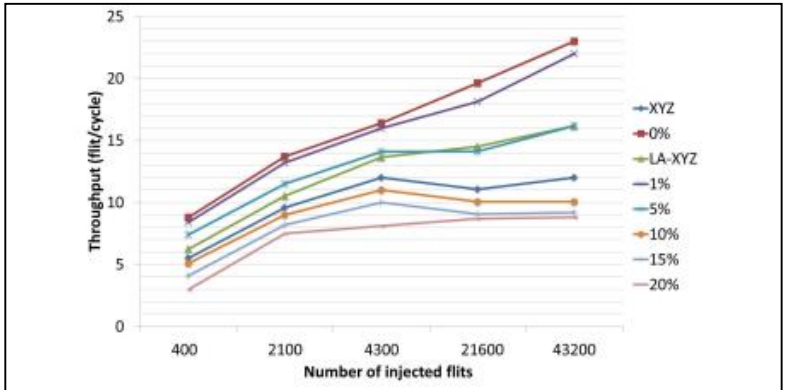


(c)

Figure 10: Latency per flit evaluation with: (a) Transpose (b) Uniform (c) 6x6 Matrix.

Table 3: Hardware complexity comparison results.

Target device	System	Area	Power (mW)			Speed (MHz)
			Static	Dynamic	Total	
FPGA	LAFT	3272	1296.92	178,09	1475.03	178.51
	LA-XYZ	3093	1264.36	169	1433.36	188.68
	XYZ	2809	1258.01	165	1423.01	194.61
Structured-ASIC	LAFT	32420	281.6	17.1	298.7	266.29
	LA-XYZ	30646	274.25	16.24	290.49	271.53
	XYZ	27822	277.6	15.83	293.43	284.93



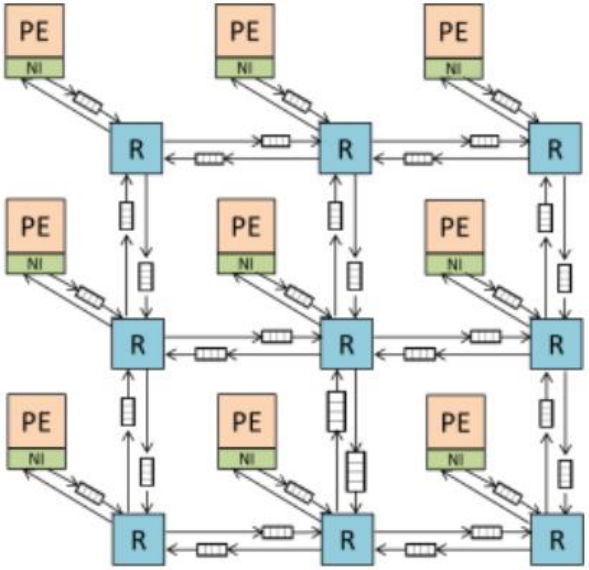
(c)

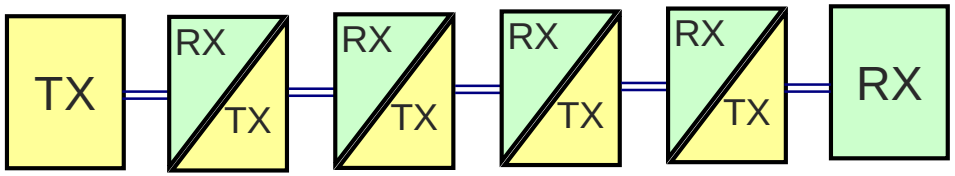
Figure 11: Throughput evaluation with: (a) Transpose (b) Uniform (c) 6x6 Matrix.

PHENIC: Hybrid Si-Photonic NoC

Replace Wires with Waveguides and Electrons with Photons!

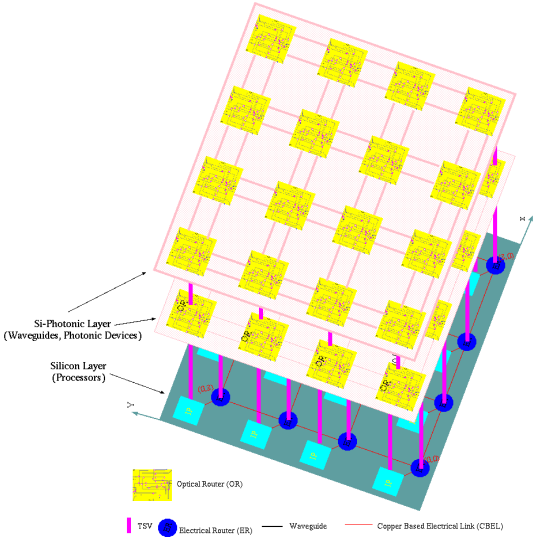
Electrical NoC

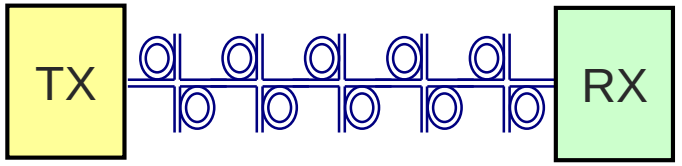




- Buffer, receive and re-transmit at every switch
- Off chip is pin-limited
- Large power/energy

Electrical-Photonic NoC



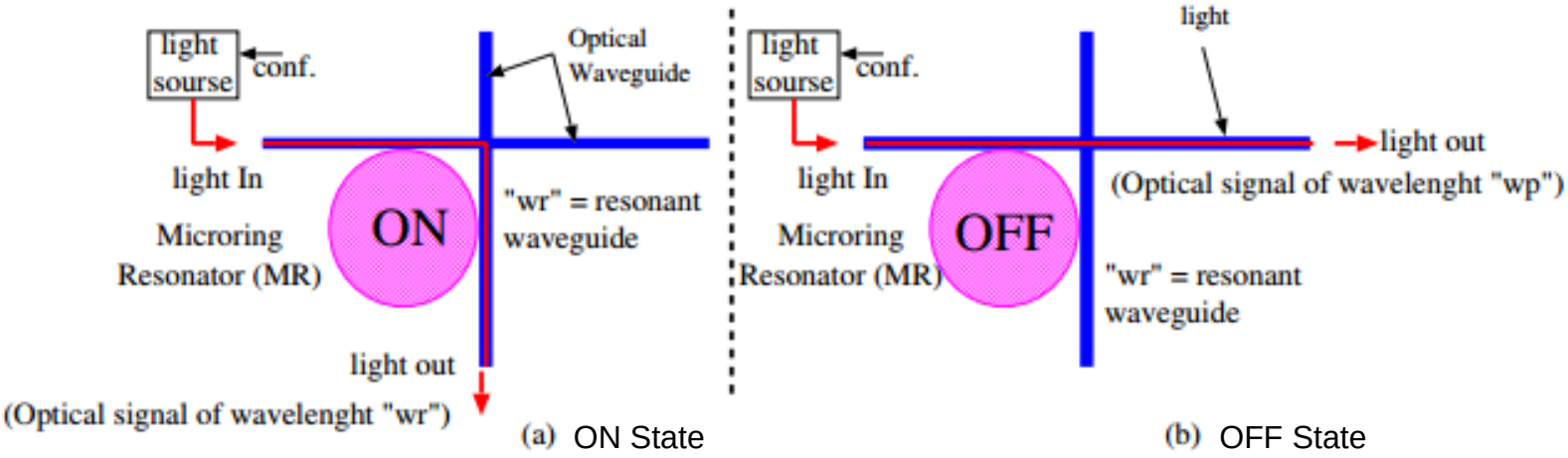


- Modulate/receive ultra-high bandwidth data stream once per communication.
- Switch routes entire multi-wavelength high BW stream
- Low power switch fabric, scalable

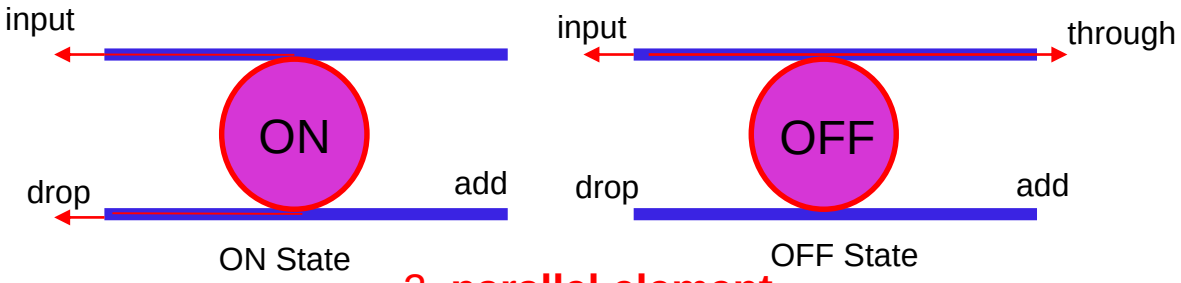
PHENIC: Hybrid Si-Photonic NoC

Replace Wires with Waveguides and Electrons with Photons!

Basic Optical Switching Element



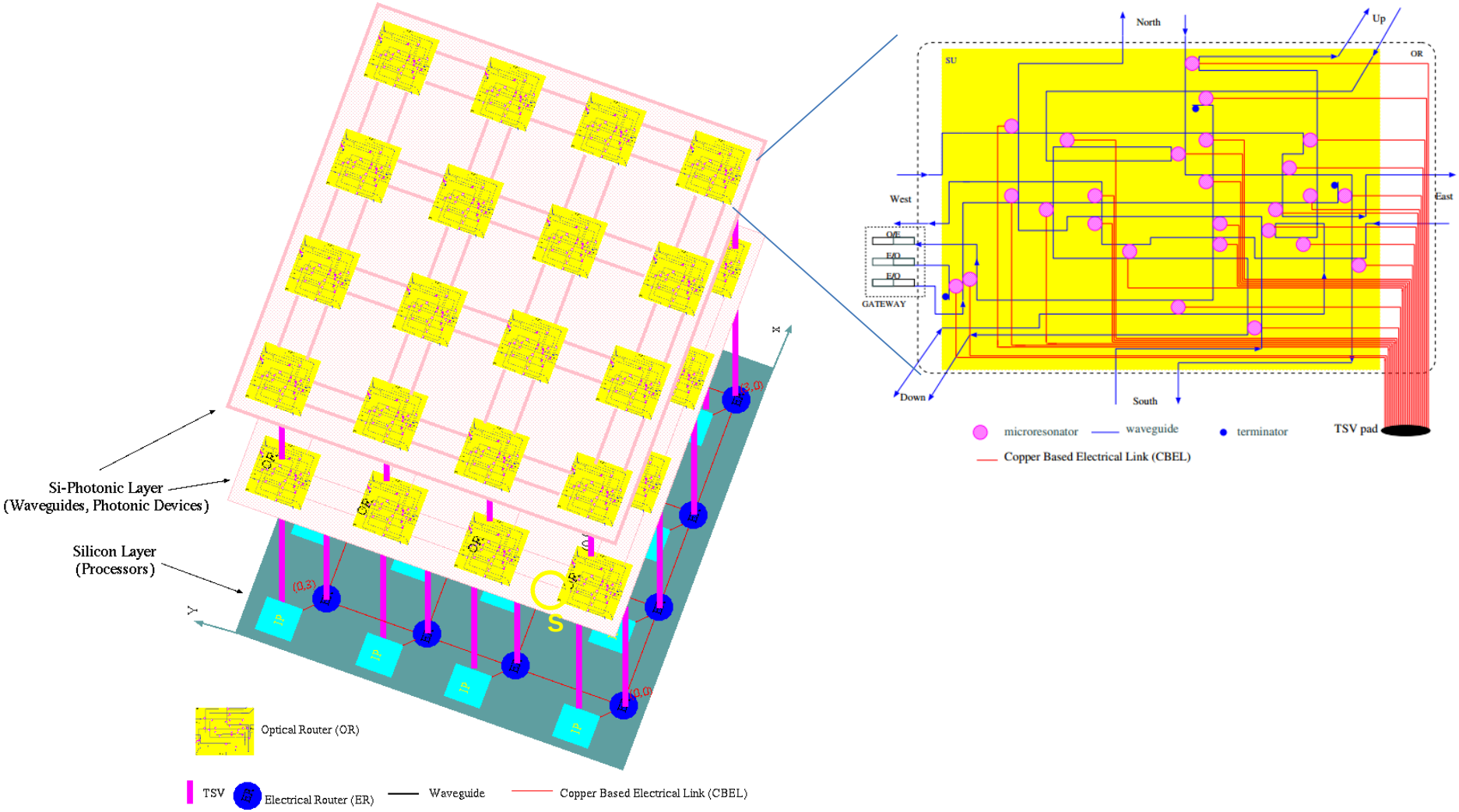
1. crossing element



2. parallel element

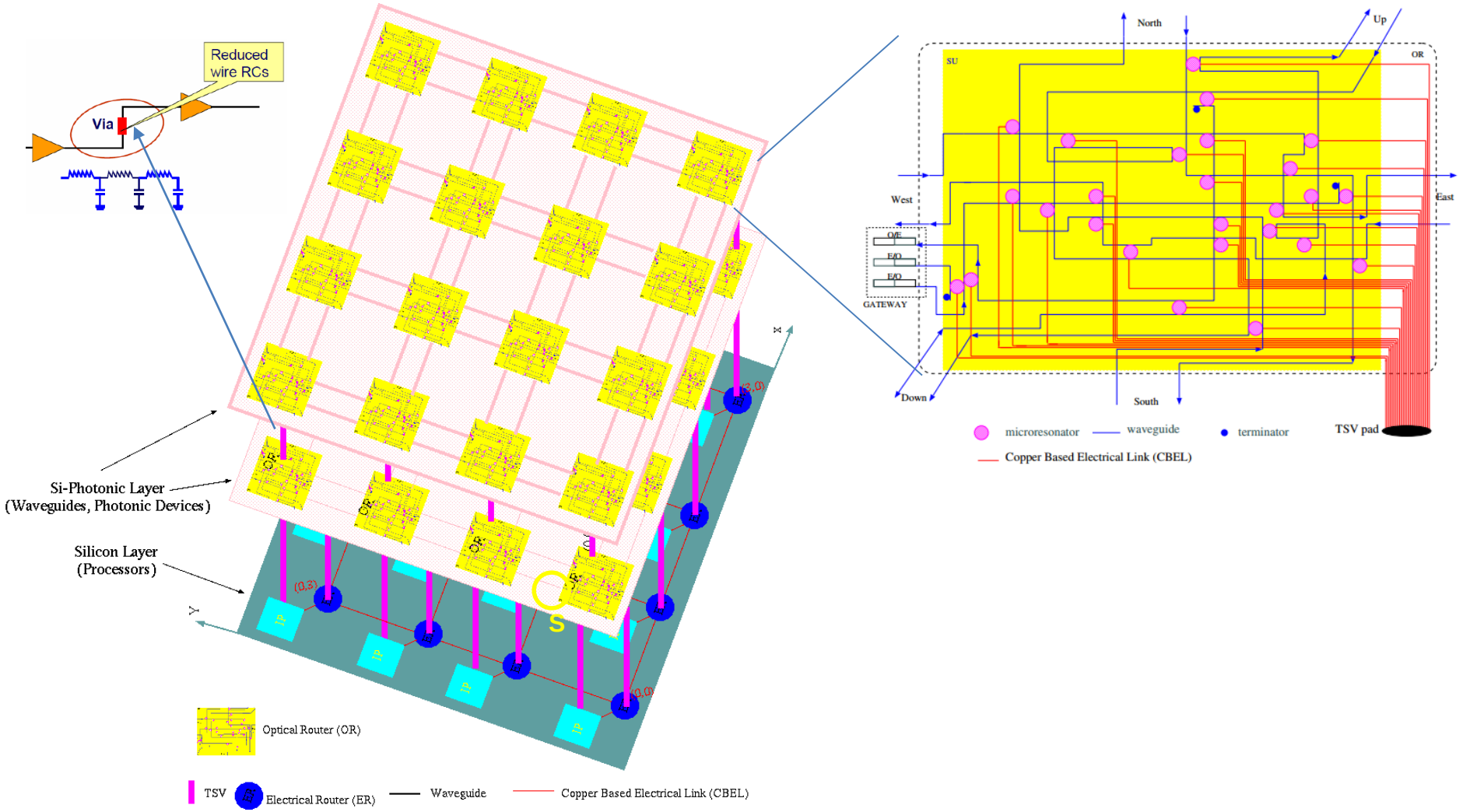
PHENIC: Hybrid Si-Photonic NoC

Replace Wires with Waveguides and Electrons with Photons!



PHENIC: Hybrid Si-Photonic NoC

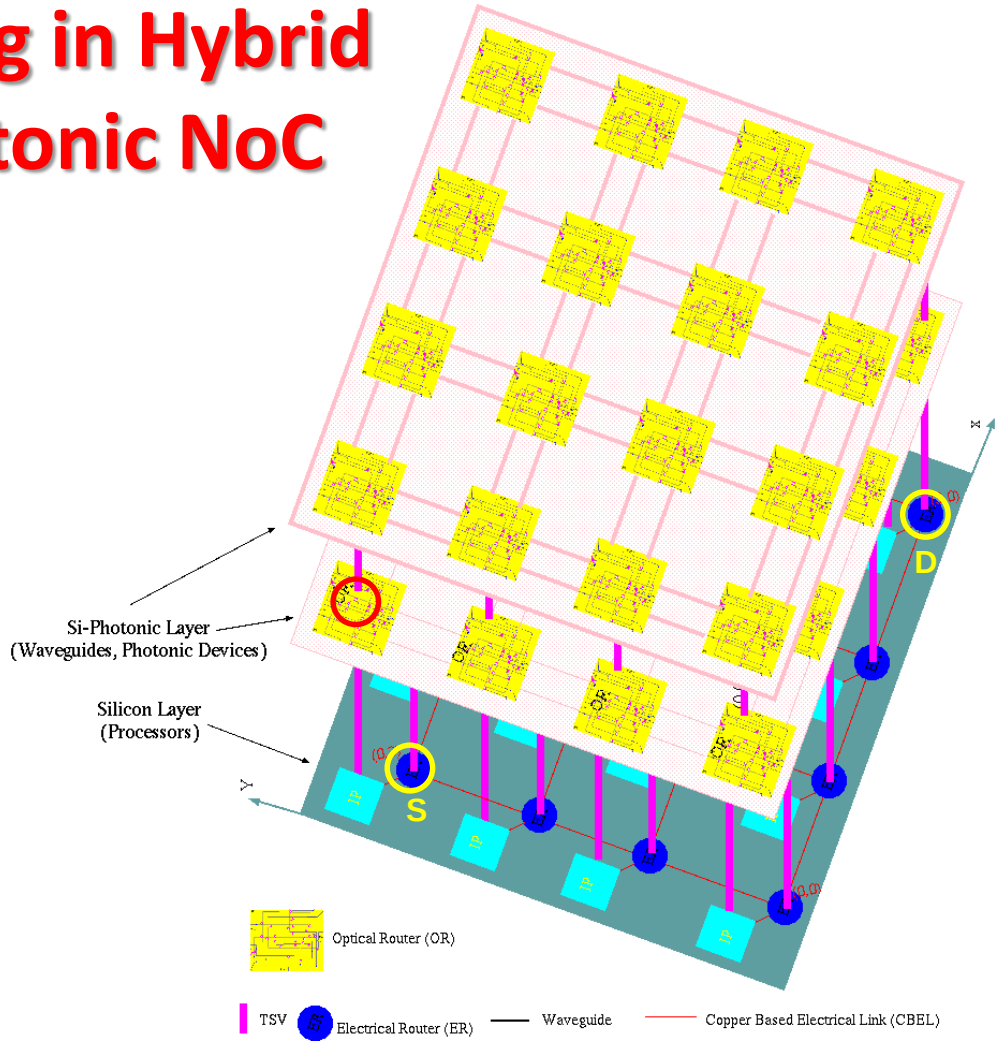
Replace Wires with Waveguides and Electrons with Photons!



PHENIC: Hybrid Si-Photonic NoC

Replace Wires with Waveguides and Electrons with Photons!

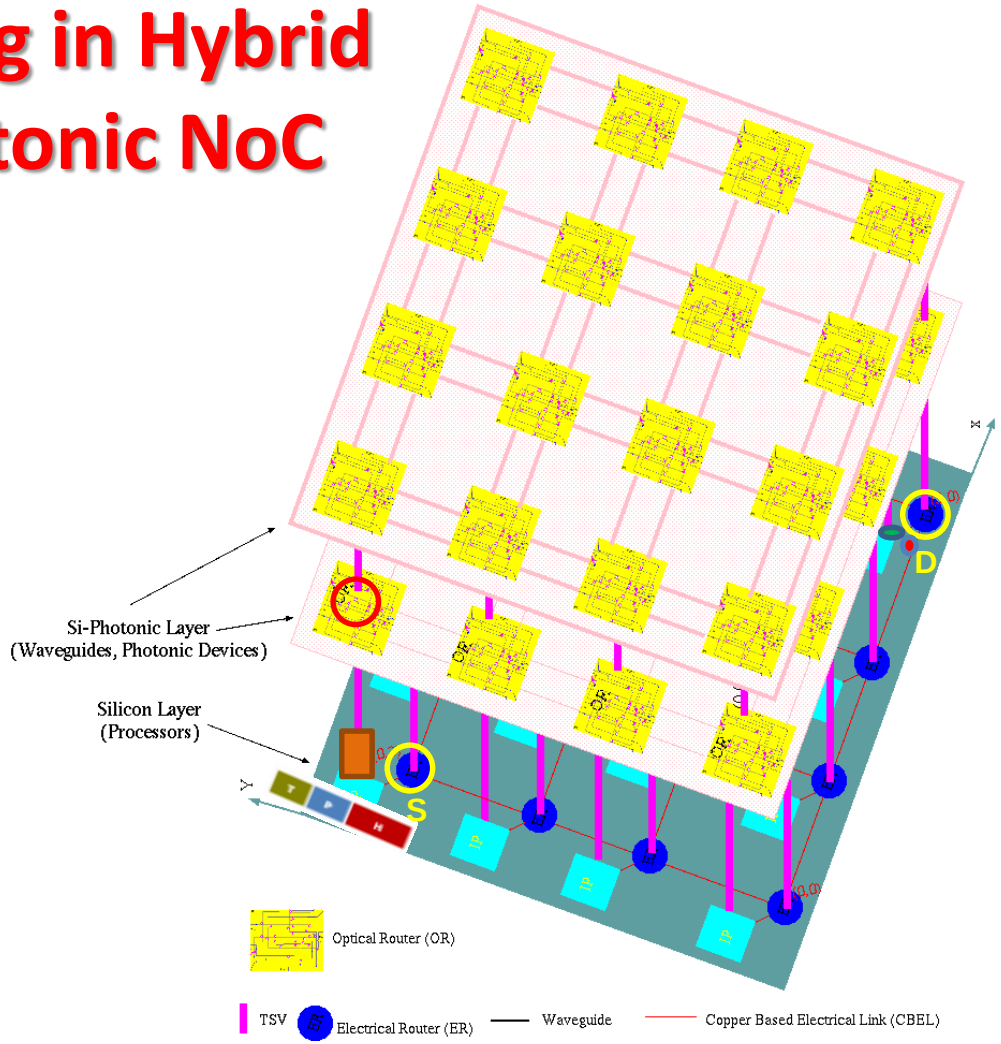
Routing in Hybrid Si-Photonic NoC



PHENIC: Hybrid Si-Photonic NoC

Replace Wires with Waveguides and Electrons with Photons!

Routing in Hybrid Si-Photonic NoC

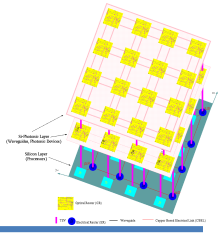


1.Reserve the path

2.ACK

3. Transmit data on the Photonic layer

4.Release (tear-down)



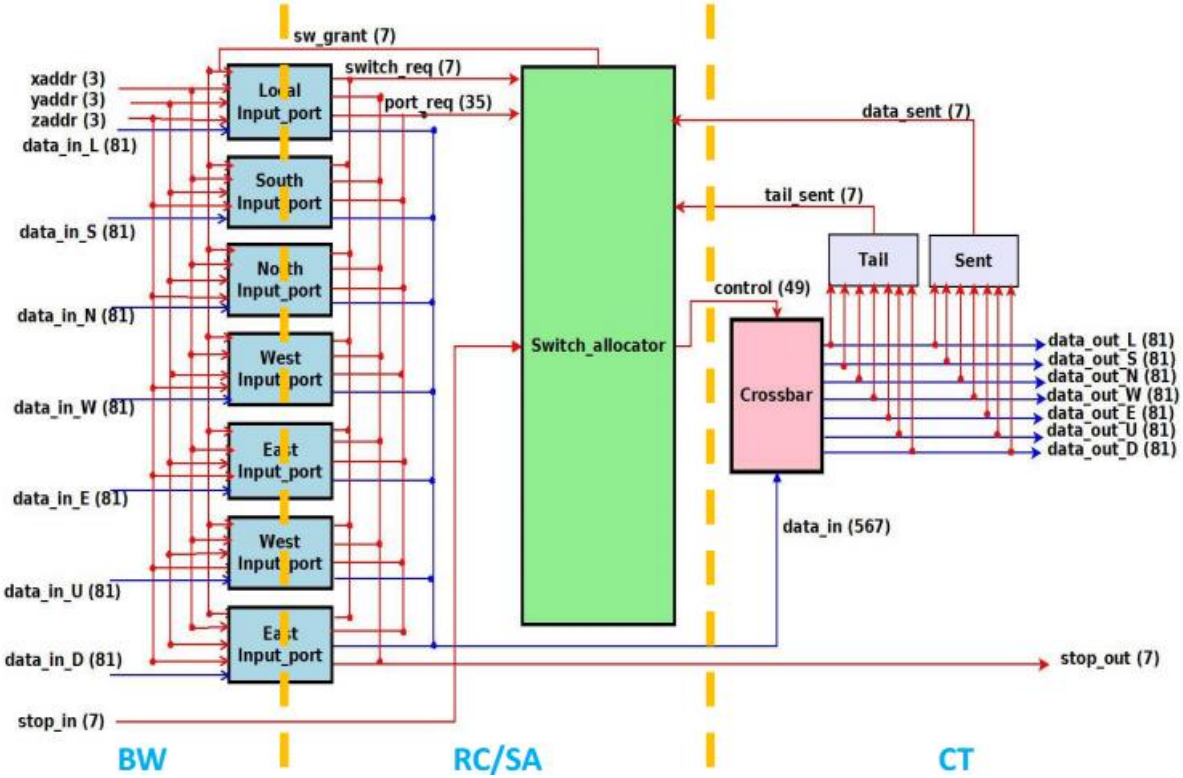
Routing in Hybrid Si-Photonic NoC

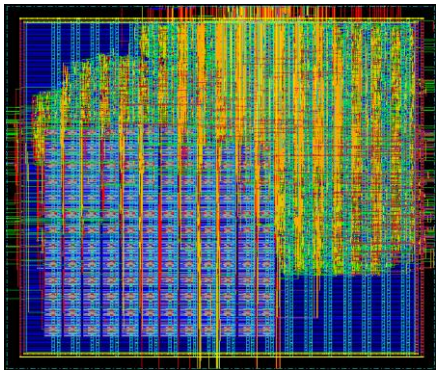
1. Reserve the path
   - ❖ A path setup message is sent by the source in the electrical network to establish a path for the optical network.
2. ACK
   - ❖ A pulse is sent back to the source node by the destination node in the optical network, and optical data can be transferred.
3. Transmit data on the Photonic layer
4. Release (tear-down)
   - ❖ Teardown message is sent by the source node in the electrical control network to release the optical circuit.

PHENIC: Hybrid Si-Photonic NoC

Replace Wires with Waveguides and Electrons with Photons!

E-Router for Path Setting and Short Messages



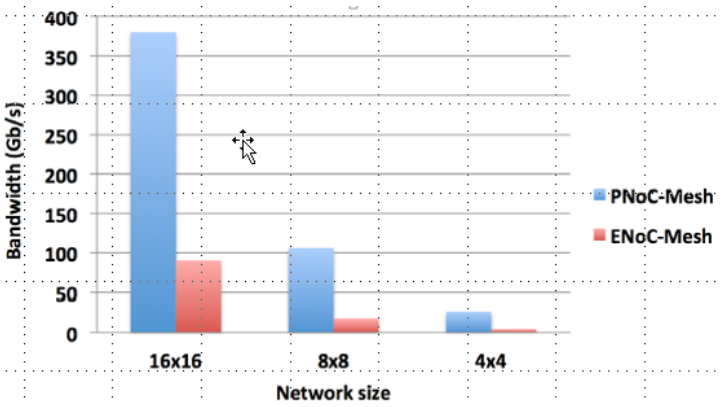


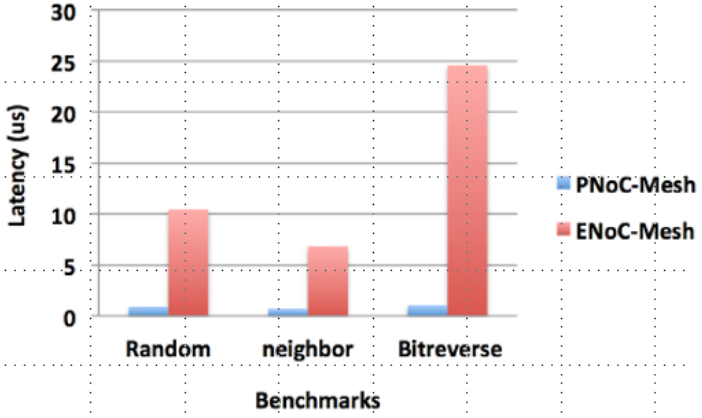
OASIS-RV2 Chip Layout

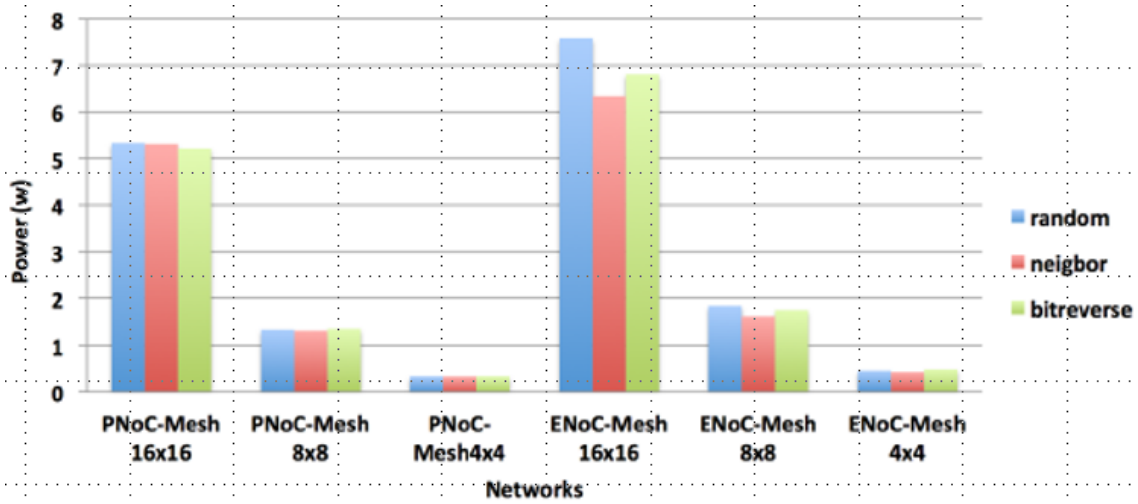
PHENIC: Hybrid Si-Photonic NoC

Replace Wires with Waveguides and Electrons with Photons!

Bandwidth, power and latency

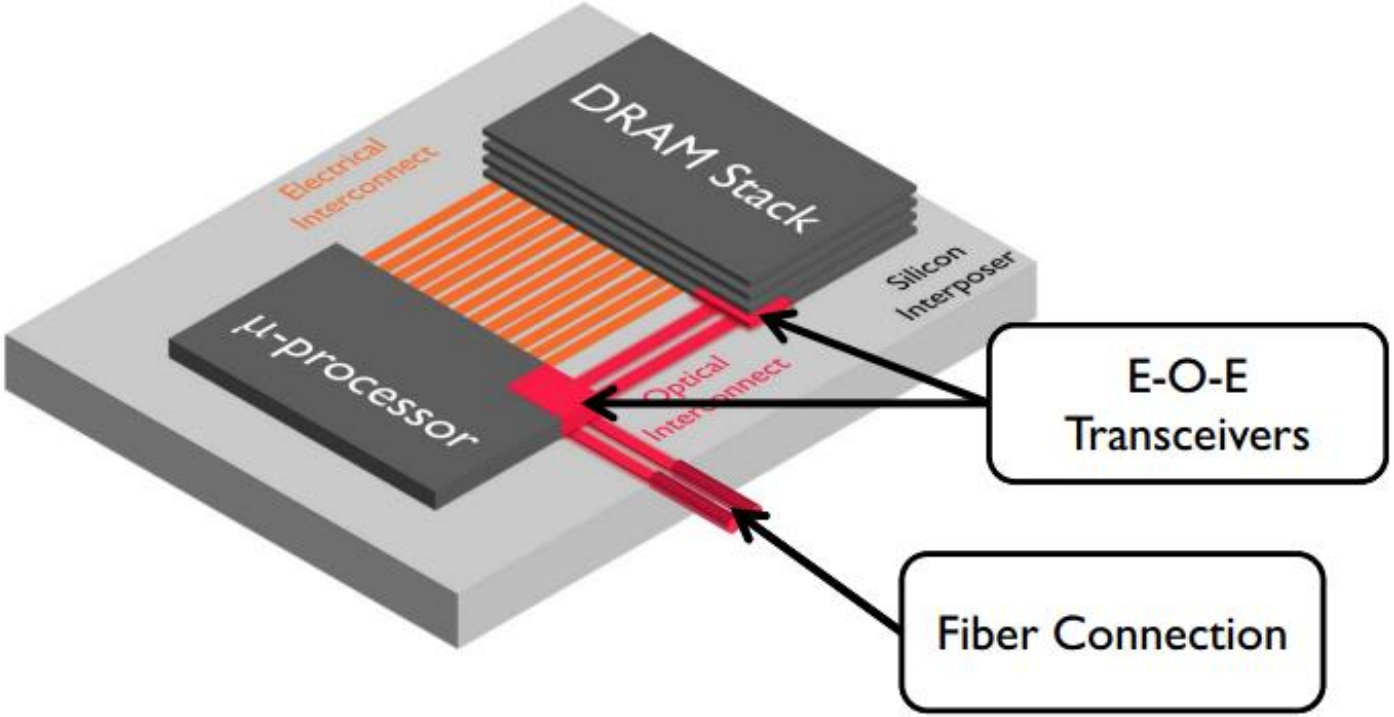






Contents

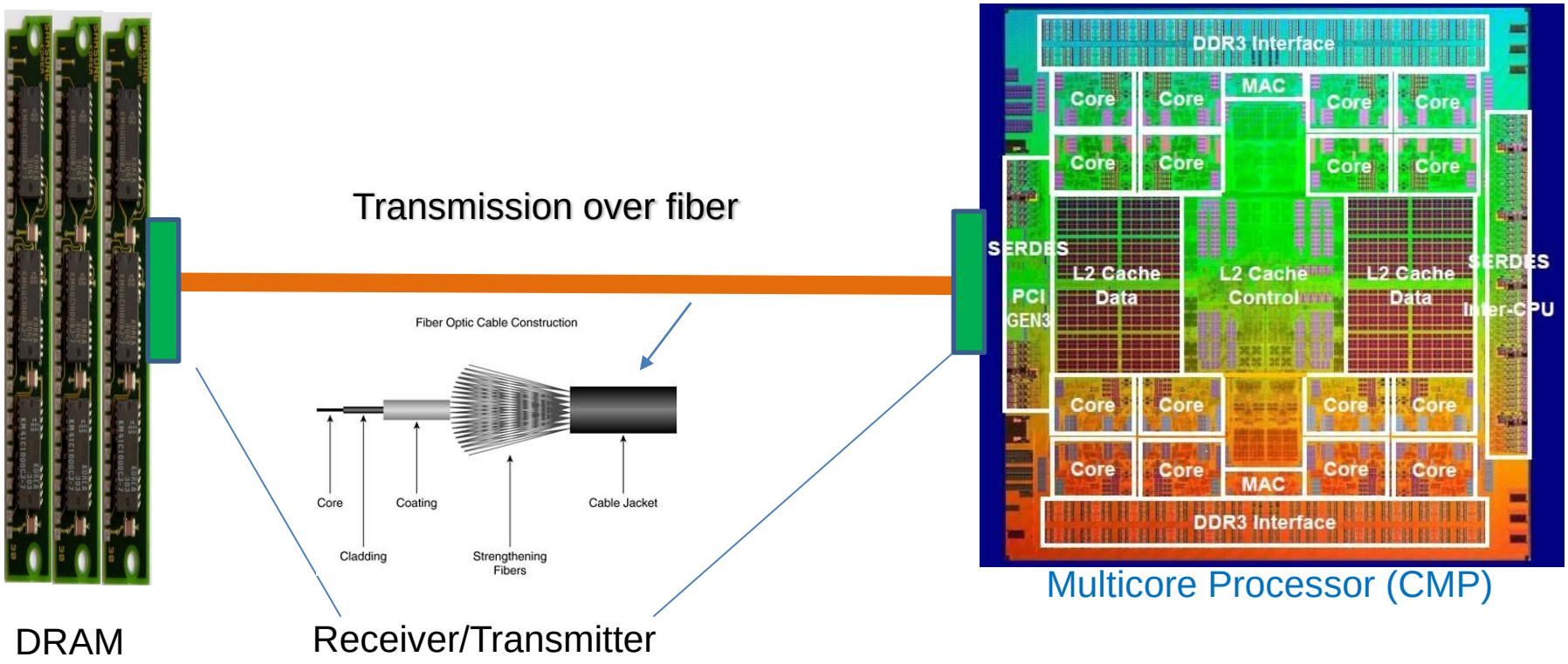
Si-Photonics in computing system today



Si-Photonics interposer

- Optical I/O's for chip-to-chip and chip-to-board links (IBM, Intel, Fujitsu)
- E-O-E transceivers for Opto-Silicon Interposer

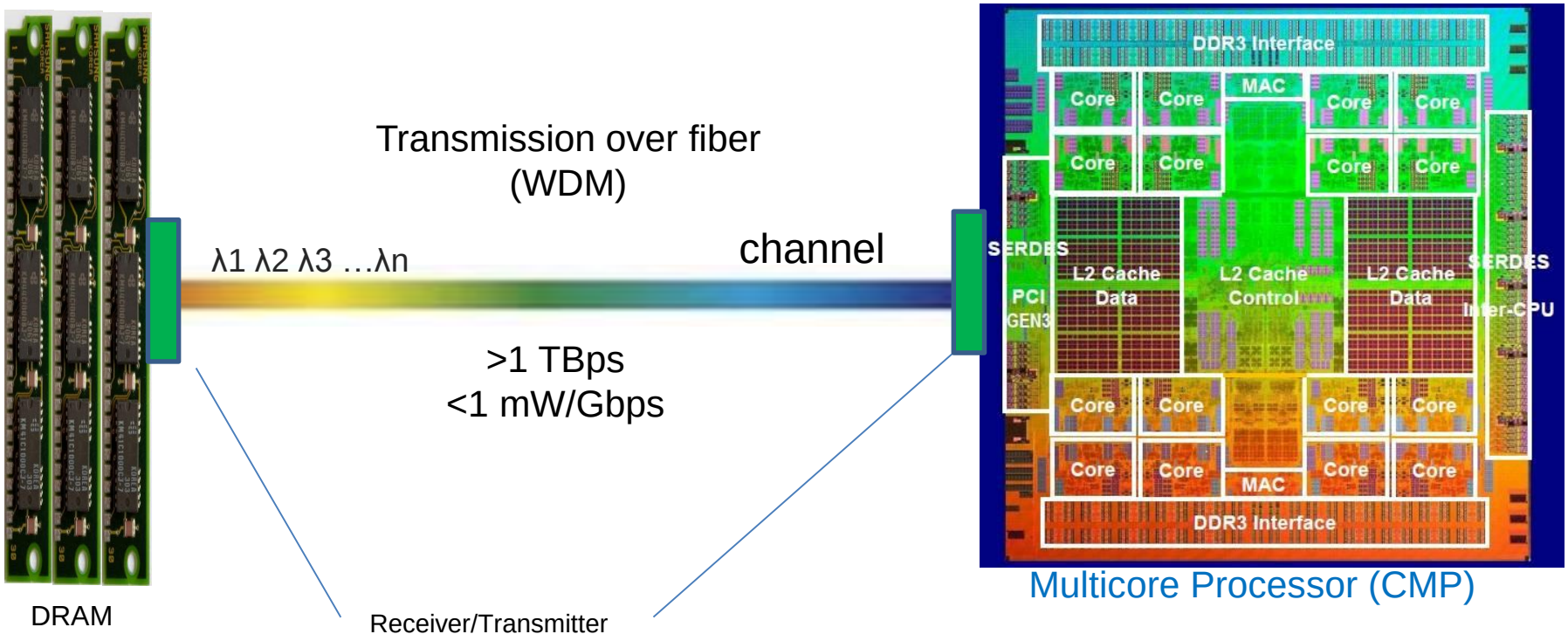
Photonics in computing system



Optical link

- Uses monolithic integration that reduces energy consumption
- Utilizes the standard bulk CMOS flow
- Cladding is used to increase the total internal reflection → reduces data loss

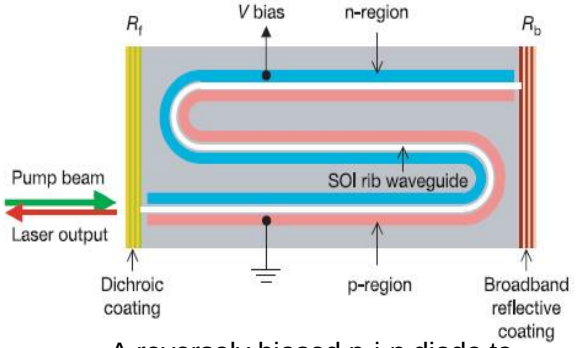
Photonics in computing system



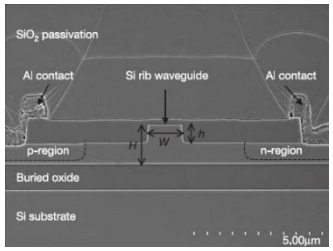
WDM, DWDM

- Supports WDM that improves bandwidth density
- DWDM can transports tens to hundreds of wavelengths per fiber.
- Integrated Tb/s optical link on a single chip is ongoing

Current Research in Photonic Components

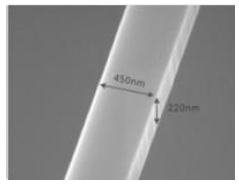


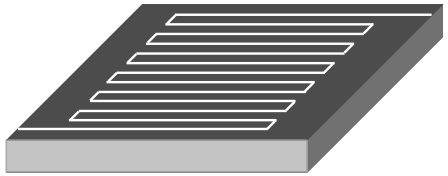
A reversely biased p-i-n diode to eliminate the TPA-induced FCA



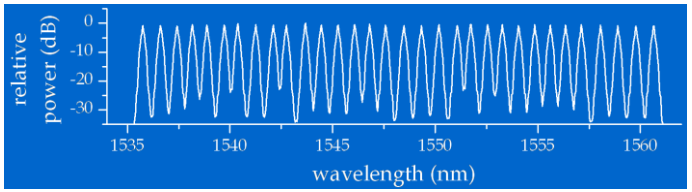
Raman Silicon Laser
Simulated Raman Scattering (SRS)

Laser

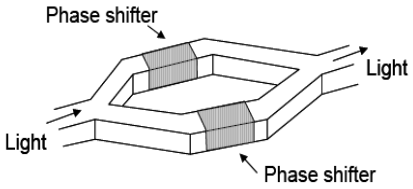


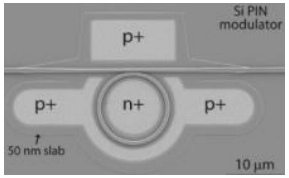


IBM/Columbia

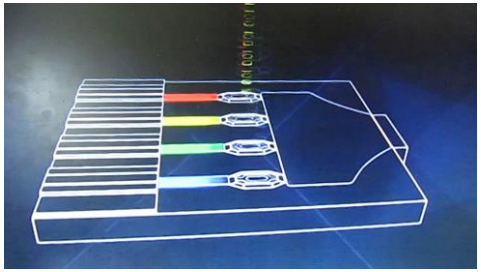


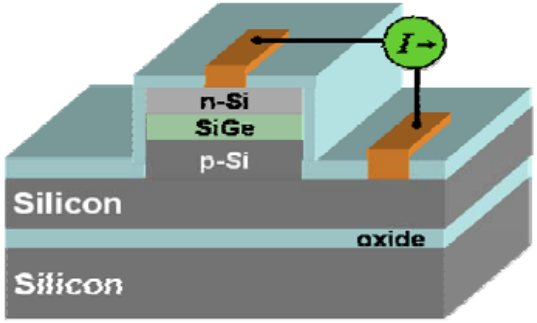
Si Wire/Waveguide

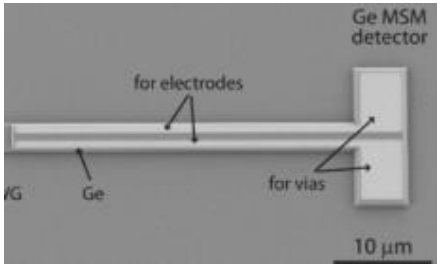




Modulator







Photodetectors

Photonic Components and Future Demands

- The necessity of low energy in optical output devices, with a ~ 10 fj/bit device energy target emerging.
  - Some Modulators and lasers meet this requirement
  - Low (few fF or less) photodetector capacitance is important
  - Very compact wavelength splitters are essential
  - Dense waveguides are also necessary on chip or on-boards for guided wave optical scheme.

Contents

Concluding remarks

- Nanophotonics will play a crucial role for on-chip interconnects
- Several technologies:
  - Si photonics, high-index contrast waveguides, photonic crystals, and plasmonics.
- Si-Photonics design approach can reduce total energy , and improve system throughput by 15-20x
  - Several approaches have been explored
  - Much more other studies should be done

References

- Achraf Ben Ahmed, A. Ben Abdallah, PHENIC: Towards Photonic 3D-Network-on-Chip Architecture for High-throughput Many-core Systems-on-Chip, IEEE Proceedings of the 14th International conference on Sciences and Techniques of Automatic control and computer engineering (STA'2013), Dec. 2013.
- A. Ben Abdallah, PHENIC: Silicon Photonic 3D-Network-on-Chip Architecture for High-performance Heterogeneous Many-core System-on-Chip>PDF, Technical Report, Ref. PTR0901A0715-2013, September 1, 2013.
- OASIS 3D-Router Hardware Physical Design, Technical Report, Adaptive Systems Laboratory, Division of Computer Engineering, University of Aizu, July 8, 2014.
- Akram Ben Ahmed, A. Ben Abdallah, Graceful Deadlock-Free Fault-Tolerant Routing Algorithm for 3D Network-on-Chip Architectures, Journal of Parallel and Distributed Computing, 2014.
- Akram Ben Ahmed, Achraf Ben Ahmed, A. Ben Abdallah, Deadlock-Recovery Support for Fault-tolerant Routing Algorithms in 3D-NoC Architectures, IEEE Proceedings of the 7th International Symposium on Embedded Multicore/Many-core SoCs (MCSoC-13), pp., 2013.
- Akram Ben Ahmed, A. Ben Abdallah, Architecture and Design of High-throughput, Low-latency and Fault Tolerant Routing Algorithm for 3D-Network-on-Chip, The Jnl. of Supercomputing, December 2013, Volume 66, Issue 3, pp 1507-1532.
- Akram Ben Ahmed, T. Ouchi, S. Miura, A. Ben Abdallah, ''Run-Time Monitoring Mechanism for Efficient Design of Application-specific NoC Architectures in Multi/Manycore Era'', '''IEEE Proc. of the 6th International Workshop on Engineering Parallel and Multicore Systems (ePaMuS2013'), July 2013.'''
- Akram Ben Ahmed, T. Ouchi, S. Miura, A. Ben Abdallah, Run-Time Monitoring Mechanism for Efficient Design of Application-specific NoC Architectures in Multi/Manycore Era, Proc. IEEE 6th International Workshop on Engineering Parallel and Multicore Systems (ePaMuS2013'), July 2013.
- Akram Ben Ahmed, A. Ben Abdallah, ''Low-overhead Routing Algorithm for 3D Network-on-Chip'', '''IEEE Proc. of the The Third International Conference on Networking and Computing (ICNC'12), pp. 23-32, 2012.
- Akram Ben Ahmed, A. Ben Abdallah, ''LA-XYZ: Low Latency, High Throughput Look-Ahead Routing Algorithm for 3D Network-on-Chip (3D-NoC) Architecture'', '''IEEE Proceedings of the 6th International Symposium on Embedded Multicore SoCs (MCSoC-12), pp. 167-174, 2012.
- Akram Ben Ahmed, A. Ben Abdallah, ''ONoC-SPL Customized Network-on-Chip (NoC) Architecture and Prototyping for Data-intensive Computation Applications'', '''IEEE Proceedings of The 4th International Conference on Awareness Science and Technology, pp. 257-262, 2012.
- Kenichi Mori,A. Ben Abdallah, OASIS Network-on-Chip Prototyping on FPGA, Master's Thesis, The University of Aizu, Feb. 2012. [
- Ben Ahmed Akram, A. Ben Abdallah,[[On the Design of a 3D Network-on-Chip for Many-core SoC, Master's Thesis, The University of Aizu, Feb. 2012.
- Shohei Miura, A. Ben Abdallah, Design of Parametrizable Network-on-Chip, '''Master's Thesis, The University of Aizu, Feb. 2012.'''
- Ryuya Okada, A. Ben Abdallah, ''Architecture and Design of Core Network Interface for Distributed Routing in OASIS NoC'', '''Graduation Thesis, The University of Aizu, Feb. 2012.'
- A. Ben Ahmed, A. Ben Abdallah, K. Kuroda, Architecture and Design of Efficient 3D Network-on-Chip (3D NoC) for Custom Multicore SoC, IEEE Proc. of the 5th International Conference on Broadband, Wireless Computing, Communication and Applications (BWCCA-2010), pp.67-73, Nov. 2010. (''best paper award'')
- Kenichi Mori, A. Ben Abdallah, OASIS Network-on-Chip Prototyping on FPGA , Master's Thesis, Graduate School of Computer Science and Engineering, The University of Aizu, Feb. 2012
- K. Mori, A. Esch, A. Ben Abdallah, K. Kuroda, Advanced Design Issues for OASIS Network-on-Chip Architecture, IEEE Proc. of the 5th International Conference on Broadband, Wireless Computing, Communication and Applications (BWCCA-2010),pp.74-79, Nov. 2010.
- T. Uesaka, OASIS NoC Topology Optimization with Short-Path Link, Technical Report, Systems Architecture Group,March 2011.
- K. Mori, A. Ben Abdallah, OASIS NoC Architecture Design in Verilog HDL, Technical Report,TR-062010-OASIS, Adaptive Systems Laboratory, the University of Aizu, June 2010.
- Shohei Miura, Abderazek Ben Abdallah, Kenichi Kuroda, PNoC: Design and Preliminary Evaluation of a Parameterizable NoC for MCSoC Generation and Design Space Exploration, The 19th Intelligent System Symposium (FAN 2009), pp.314-317, Sep.2009.
- Kenichi Mori, Abderazek Ben Abdallah, Kenichi Kuroda, ''Design and Evaluation of a Complexity Effective Network-on-Chip Architecture on FPGA'', The 19th Intelligent System Symposium (FAN 2009), pp.318-321, Sep. 2009.
- A. Ben Abdallah, T. Yoshinaga and M. Sowa, Mathematical Model for Multiobjective Synthesis of NoC Architectures, IEEE Proc. of the 36th International Conference on Parallel Processing, Sept. 4-8, 2007,

References

World Scientific
Connecting Great Minds

Springer

springer.com

Multicore Systems-on-chip: Practical Hardware/Software Design Issues
Hardcover – August 6, 2010

2013, XXVI, 273 p. 196 illus., 79 illus. in

A. Ben Abdallah

Multicore Systems On-Chip: Practical Software/Hardware Design

Series: Atlantis Ambient and Pervasive Intelligence, Vol. 7

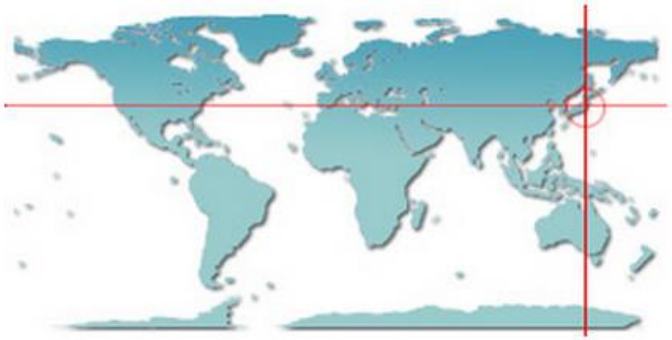
- Provides practical hardware/software design techniques for Multicore Systems-on-Chip
- Provides a real case study in Multicore Systems-on-Chip design
- Provides interaction between the software and hardware in Multicore Systems-on-Chip
- Provides detailed overview of various existing Multicore SoCs

Adaptive Systems Laboratory
http://aslweb.u-aizu.ac.jp/wiki/index.php?Adaptive%20Systems%20Laboratory

http://aslweb.u-aizu.ac.jp/

Thank you!

Abderazek Ben Abdallah
benab@u-aizu.ac.jp



to Advance Knowledge for Humanity