Photonic Networks for Data Centres and High Performance Computing

Philip Watts

Department of Electronic Engineering, UCL

Yury Audzevich, Nick Barrow-Williams,
Robert Mullins, Simon Moore, Andrew Moore

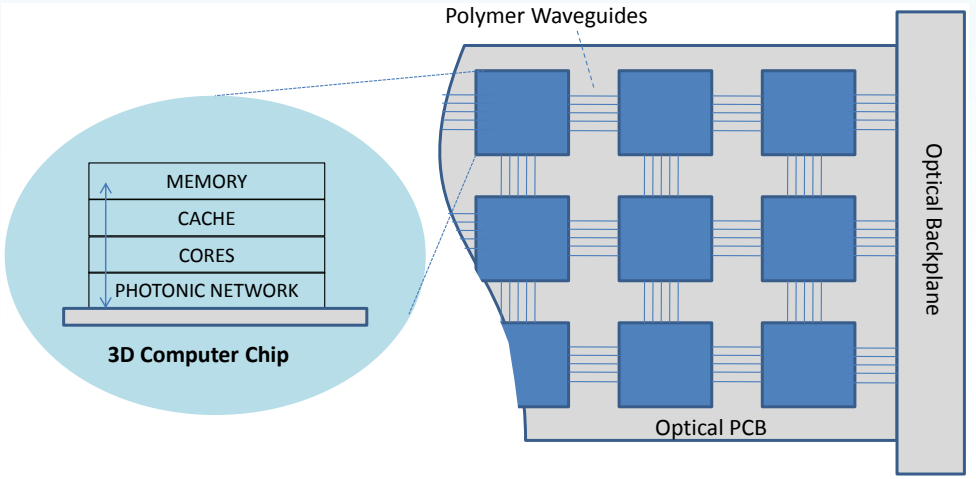
Computer Laboratory, University of Cambridge

Motivation

- Future digital services are reliant on power efficient computers
  - Data centres, large servers, scientific computing
- Power management is now the primary issue
  - 200 W/chip thermal limit reached
  - Interconnect accounts for majority of power consumption
    - Transistors scale, wires don't
- Can't continue to exploit Moore's Law without dealing with interconnect power issue
- Many developments are addressing this issue:
  - 3D stacked memory
  - Non-volatile 'unified' memory
  - Photonic interconnect

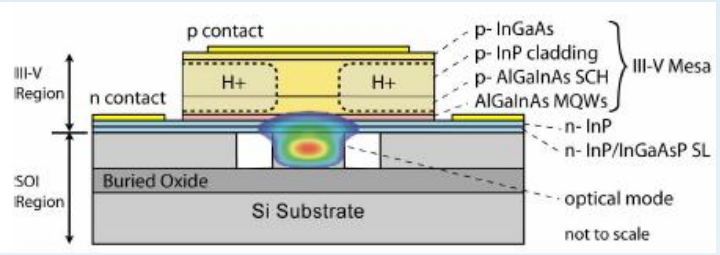
Power of Computation vs. Communication

	130nm CMOS (2002)	45nm CMOS (2008)	Increase
Transfer 32b across chip	20 computations	57 computations	+285%
Transfer 32b off-chip	260 computations	1300 computations	+500%

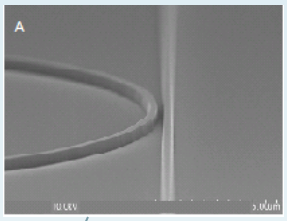


Components of a Photonic C2C Interconnect

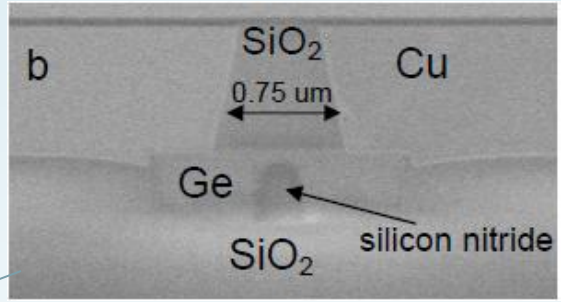
Hybrid Si Laser (UCSB/Intel)

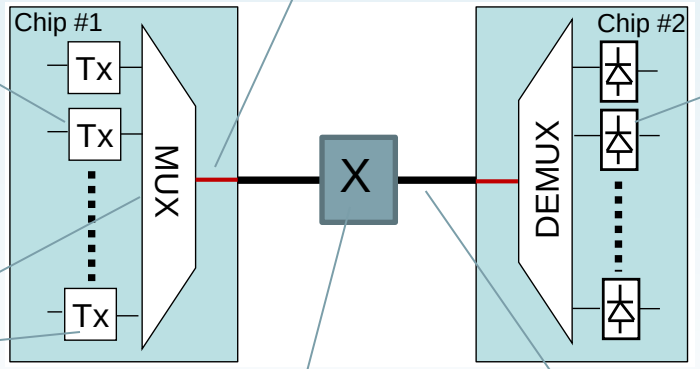


Si/SiO_2 Waveguides
400 x 200 nm, Bend Radius ≈ few µm

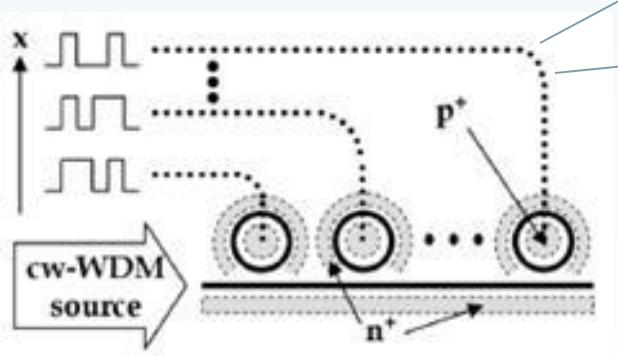


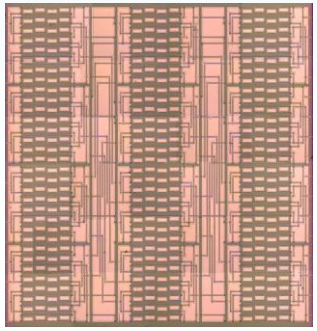
Germanium Photodiodes integrated with Silicon waveguides 3 x 1 µm (Intel)



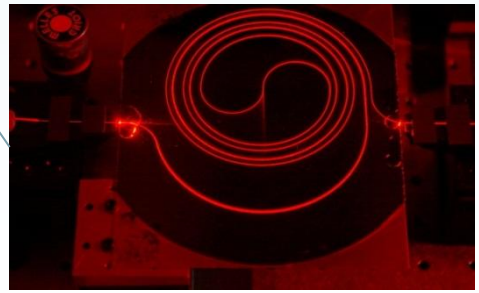


Silicon ring resonators can modulate, filter and switch





16 x 16 integrated SOA switch (Cambridge/Eindhoven)

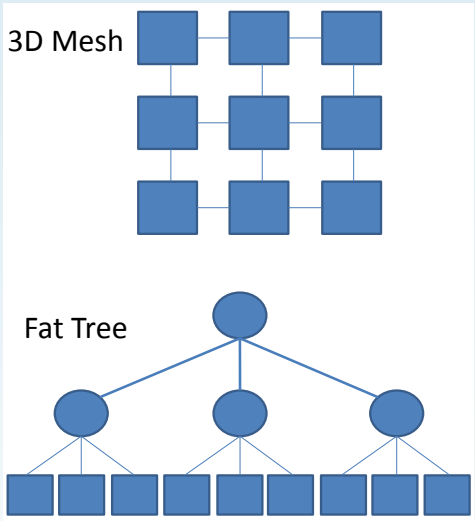


Polymer waveguides in standard FR4 PCB (Cambridge)

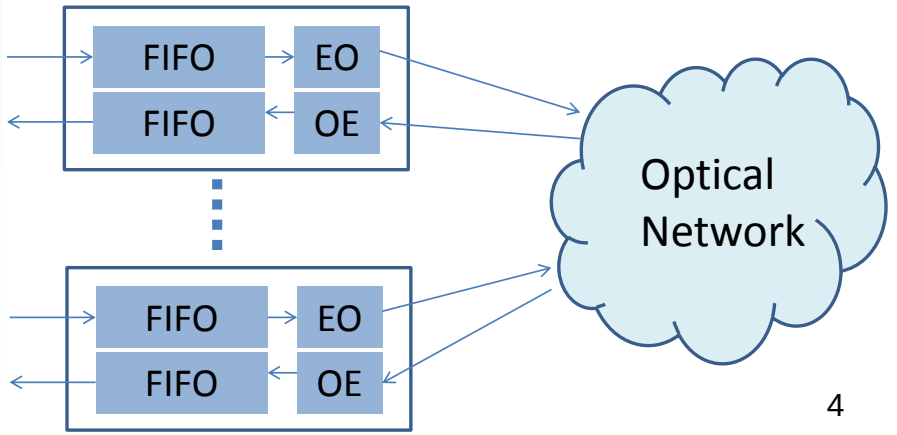
Characteristics of Photonic Networks

- Photonics offers very high bandwidth (WDM) over long distances with lower power
- Photonics favours circuit switching
  - No viable optical buffer exists
  - End-to-end paths required
- Computer architectures favours packet switching
  - Message sizes in Shared Memory systems ≥ 64-bits
- Design implications of complete computer systems with photonic interconnect are not clear

Traditional Electronic Networks



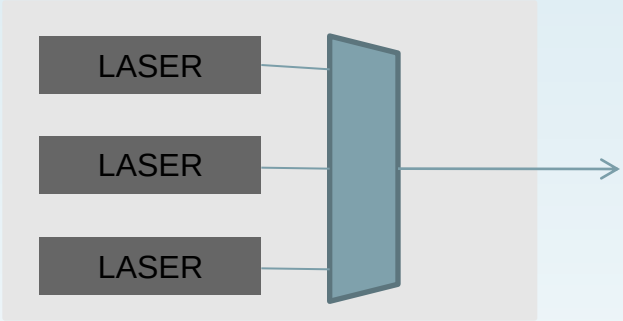
Edge Buffered Photonic Network



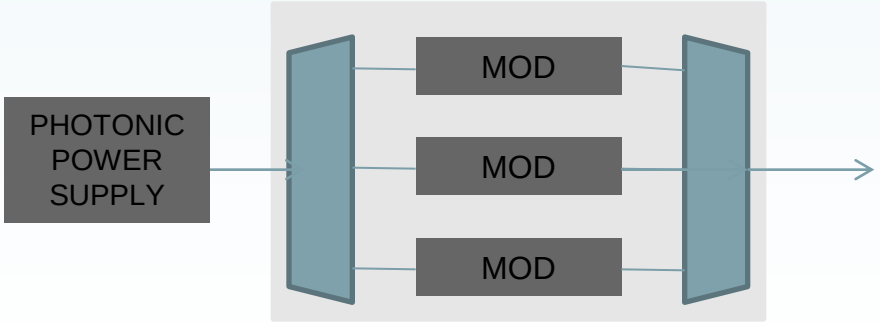
On-chip Lasers or Off-chip Power Supply

- On chip laser approach
  - Avoid losses in distribution and getting light onto chip
  - Can be rapidly power gated (≈ns)
- Photonic power supply (PPS) approach
  - Only compact and low drive power components need to be on-chip
  - Single optical power supply can be used for multiple devices
- We can't store photons – any generated light is dissipated mainly on-chip
- Unlikely that efficiency of optical power supply will be >20%
  - Efficiency of electronic power supply can be >80% and is naturally dynamic

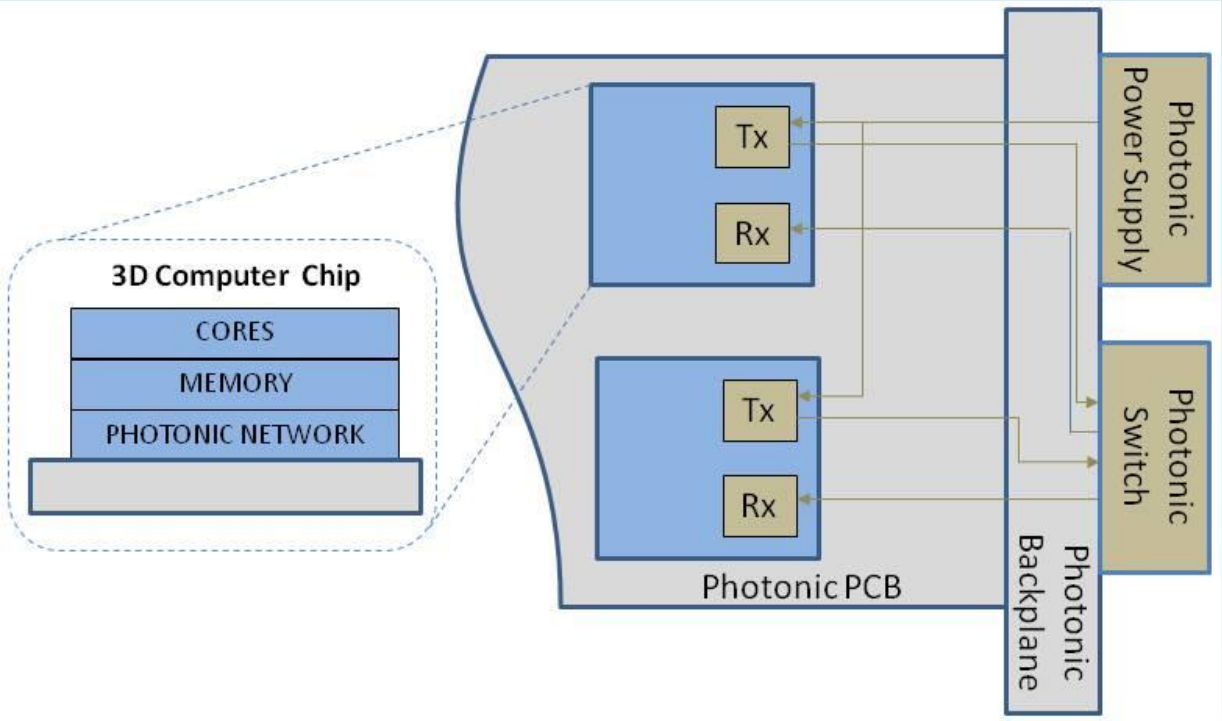
On-chip laser approach



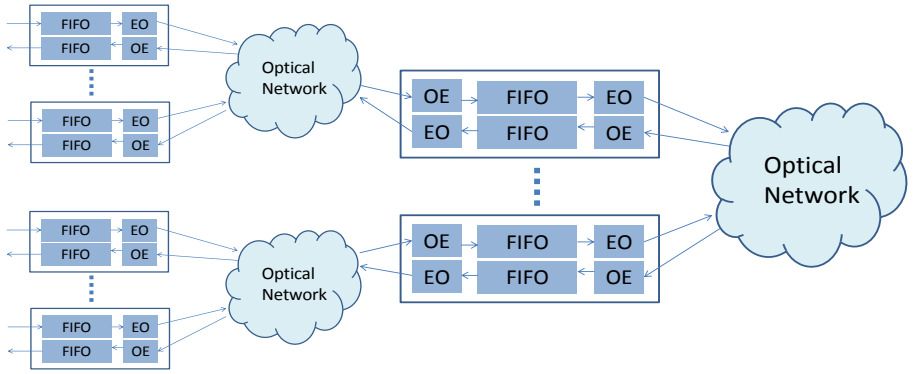
Photonic Power Supply approach



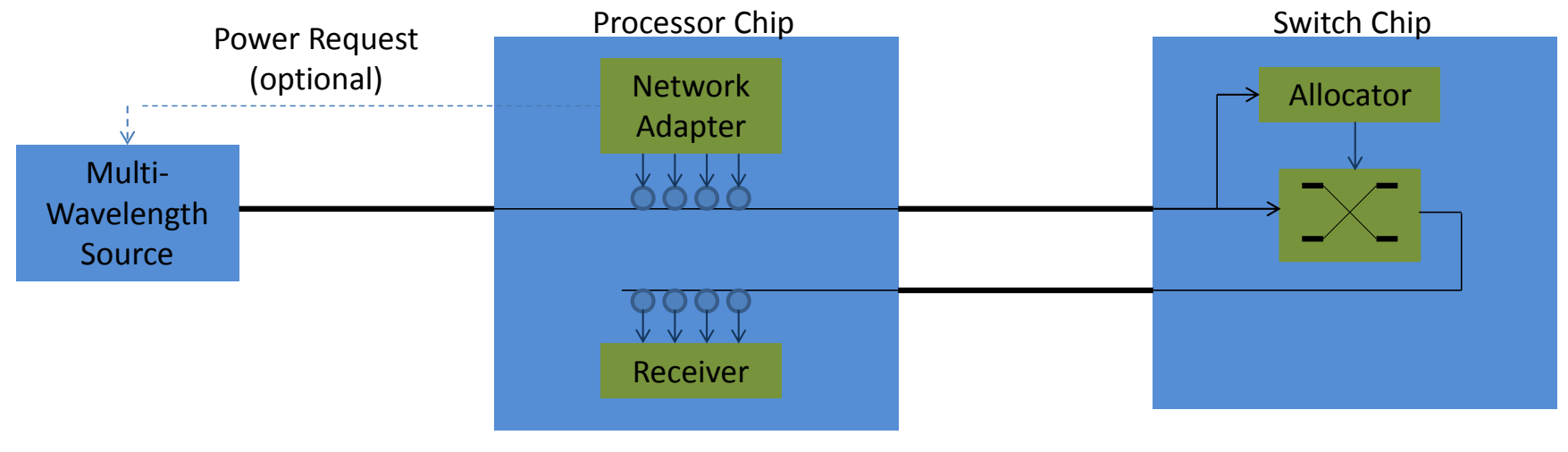
Photonic Networks for Data Centres - Vision



- Scalability of the all-optical network is limited to around 64 ports, rack network
  - Optical power and attenuation
  - Latency and complexity of Arbitration/Contention resolution
- Expandable using OEO sections



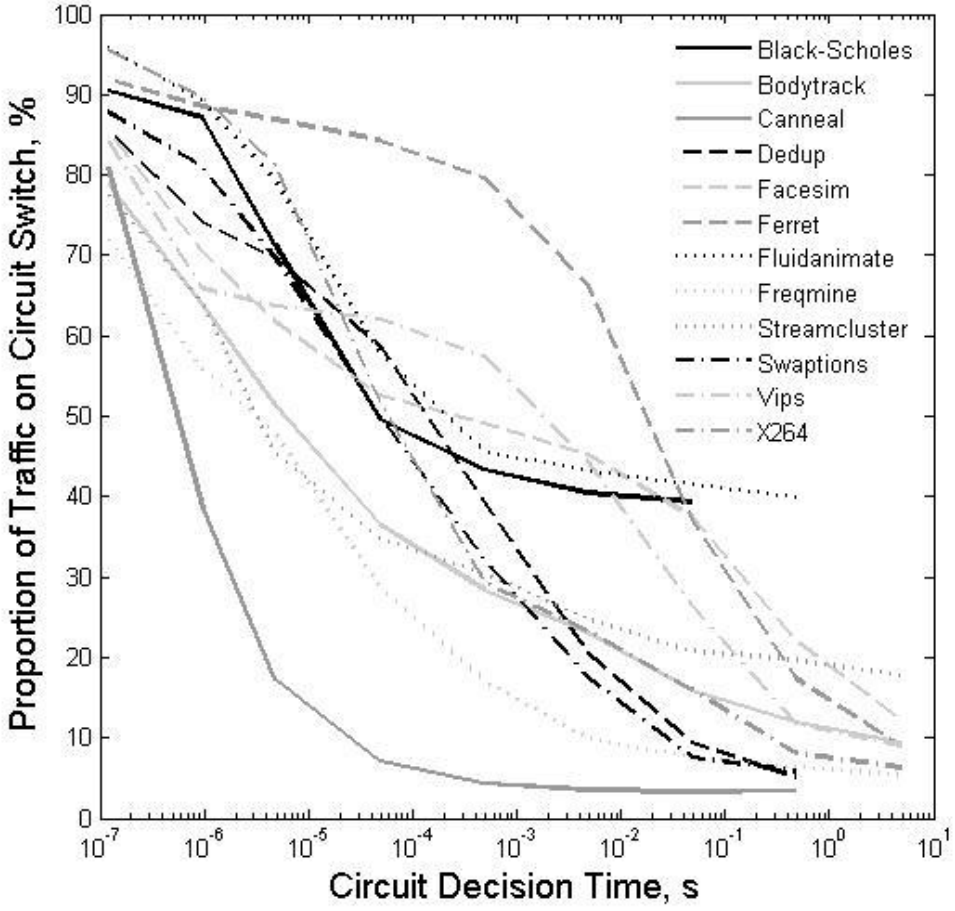
Methodology



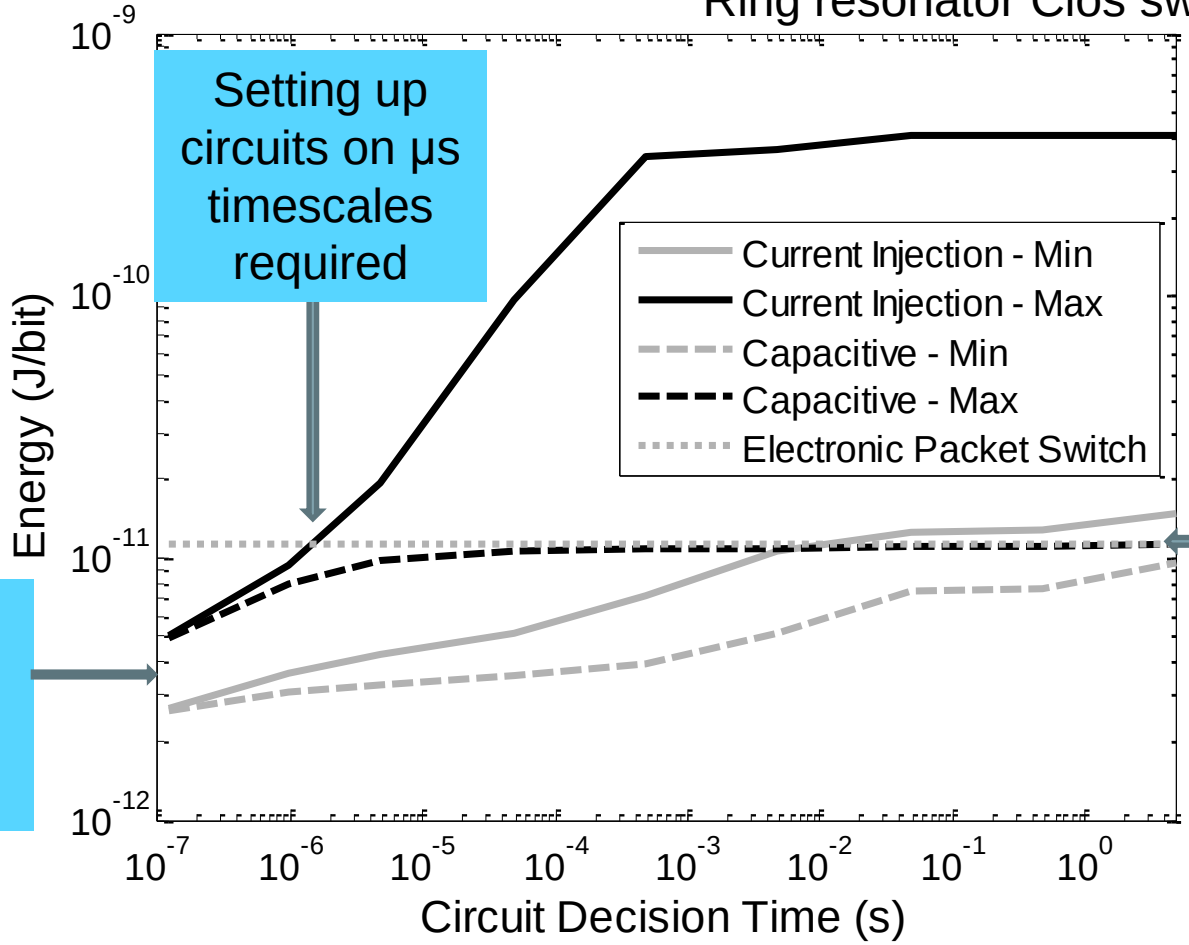
- Modelled the power consumption of the photonic path using the photonic power supply approach
- Created SystemVerilog models of the network control circuits and synthesized using a low leakage 45 nm CMOS process
- Evaluated the power consumption of circuit switched and time slotted networks covering a single rack
  - How much power is disipated on-chip?
  - What can be power gated?

Networks: Circuit Switching

- Electronic Network is required for residual flows
- Analysis of the traffic patterns generated running the PARSEC benchmark suite on simulated 32-core x86 system running Linux
- Observe the effect of changing the circuit decision time
  - Circuit decisions are ideal (a priori knowledge of traffic)

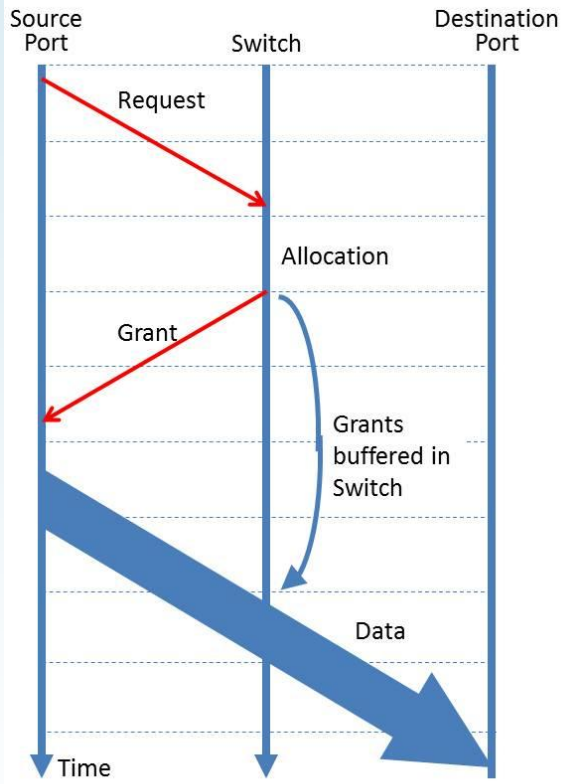


Energy per bit – Circuit Switching

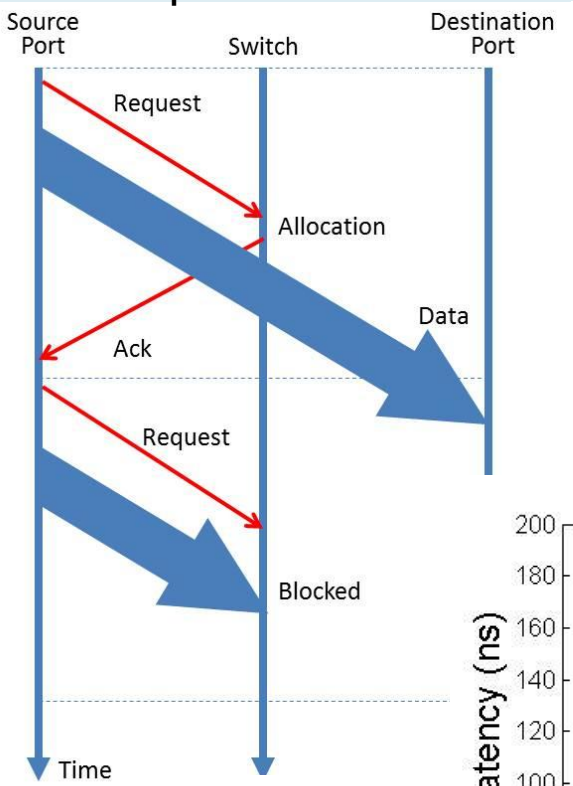


Time Slotted Networks

Scheduled



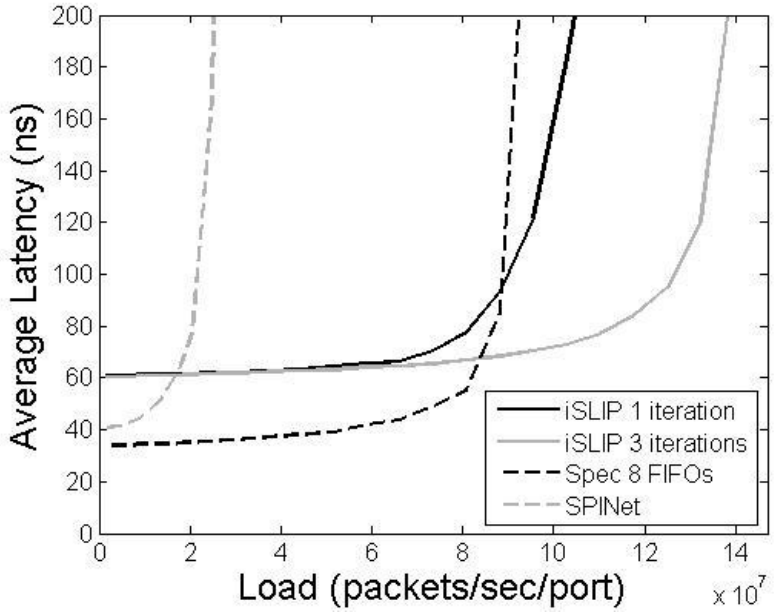
Speculative



Time slots are 10 ns, sized for transmission of a single 32B cache line

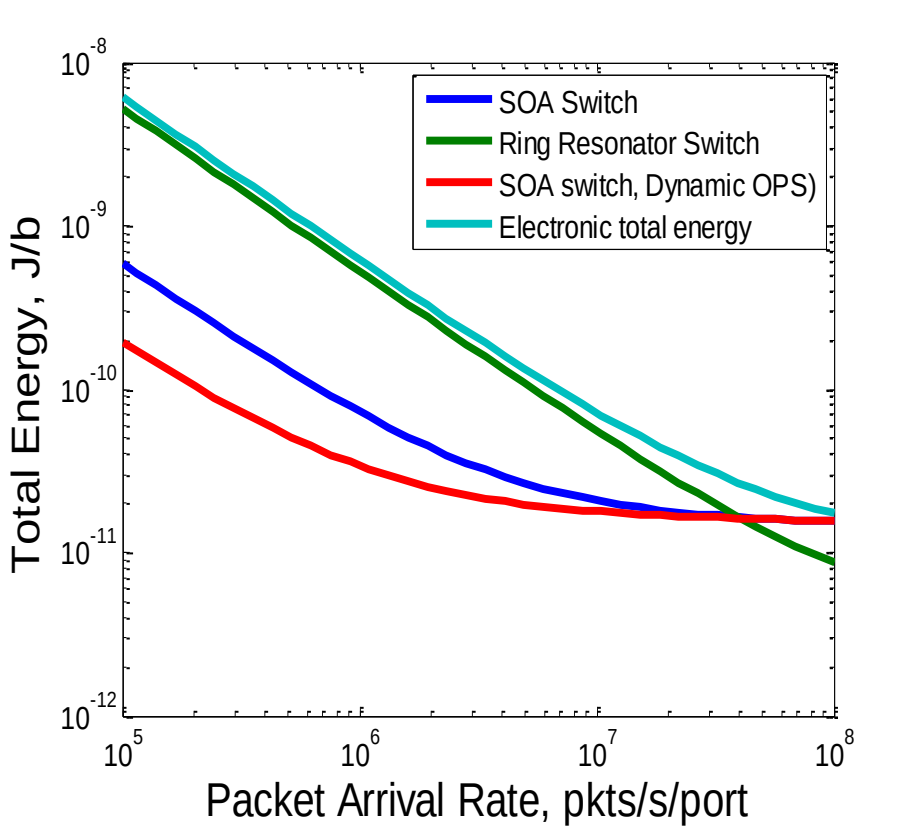
Speculative does not allow power gating of PPS

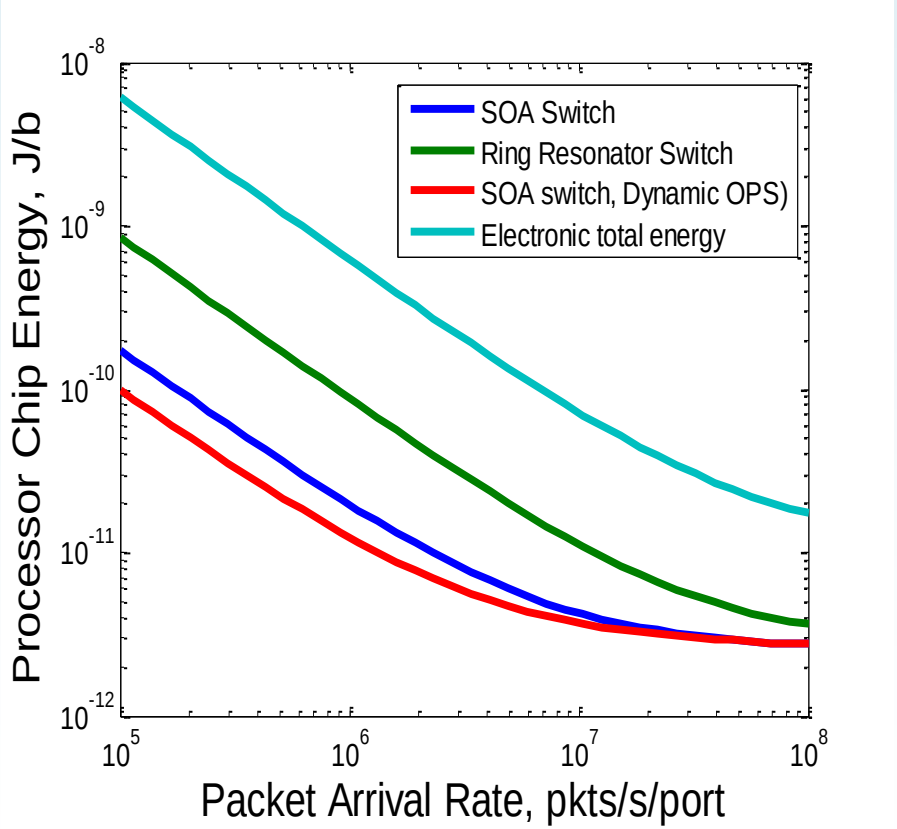
Intra-Rack Network Latency



Time Slotted Network

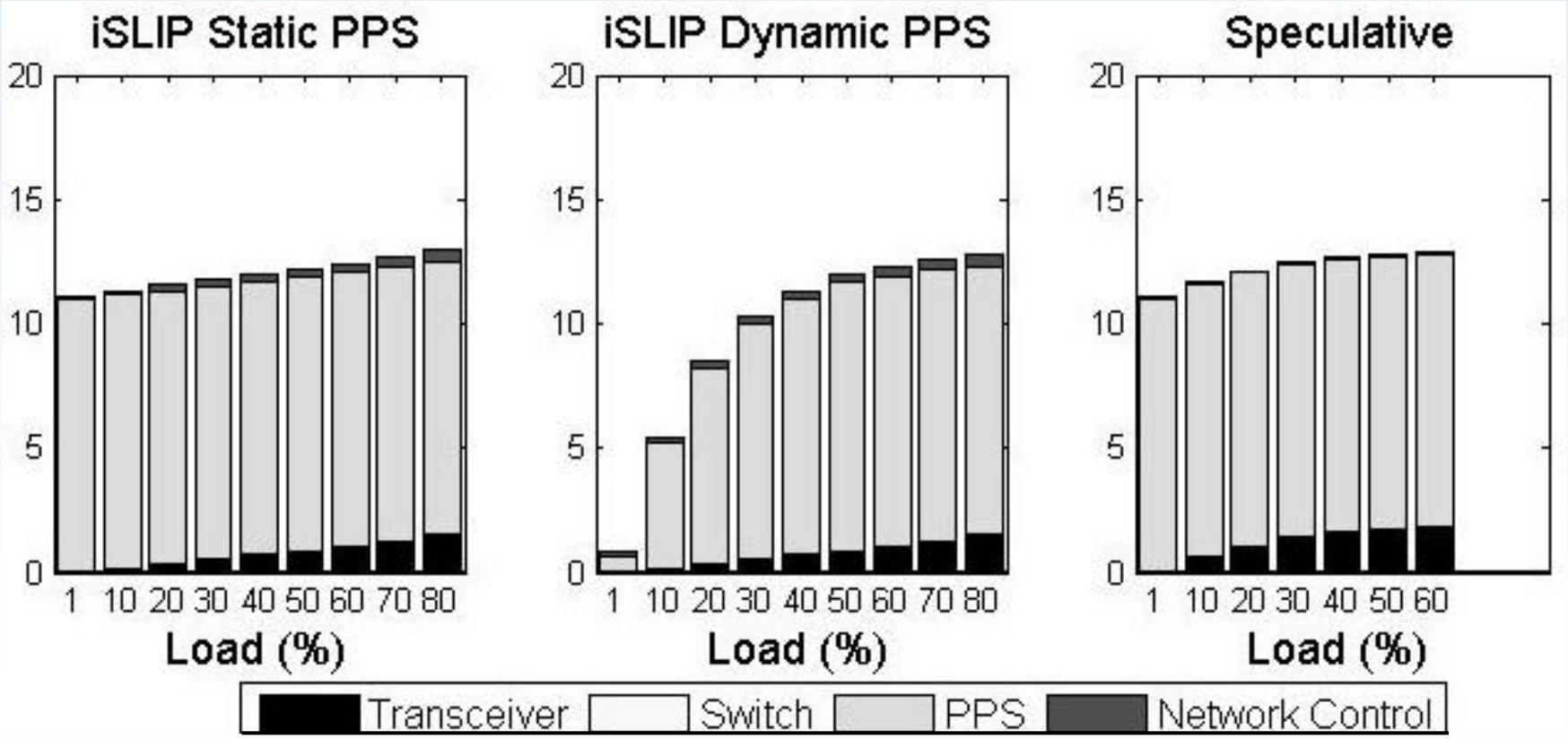
Time slotted networks can be > order of magnitude lower power than electronic mesh network





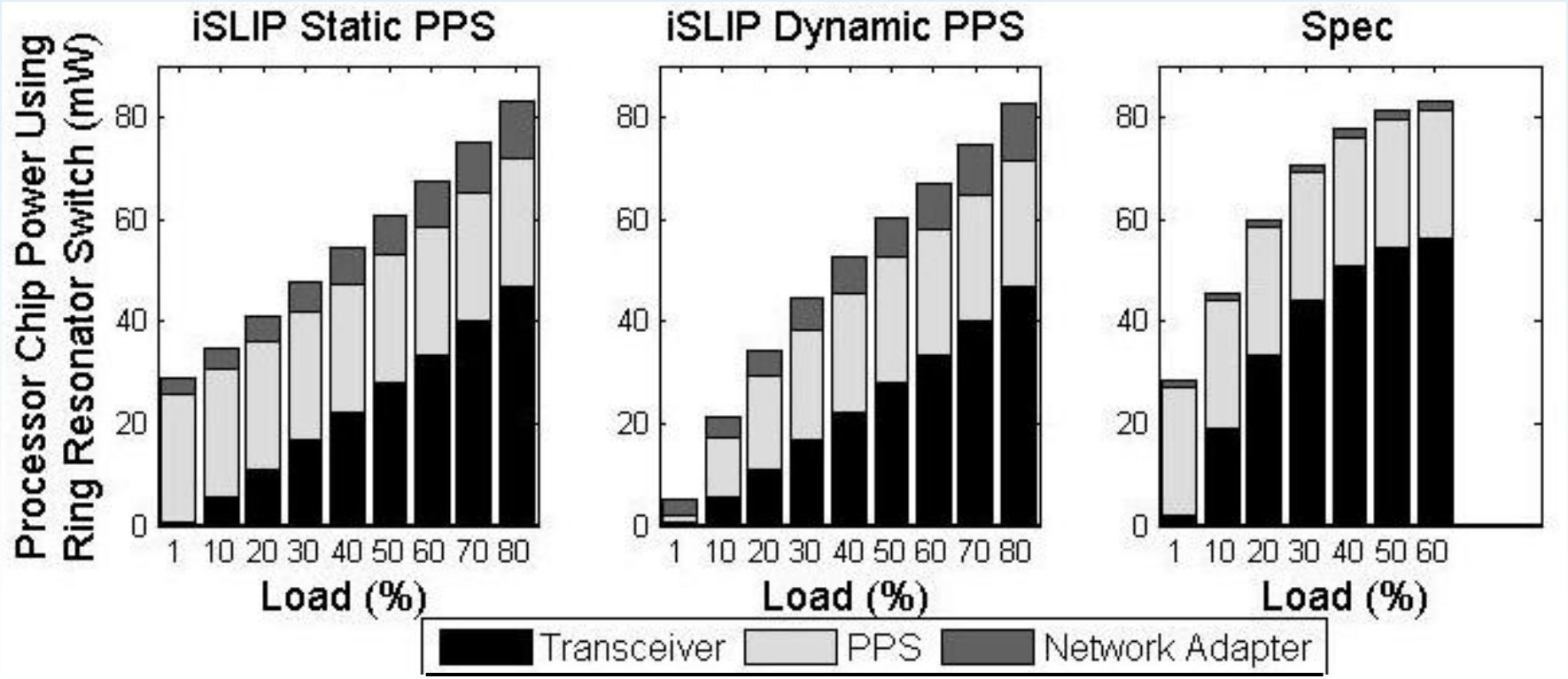
Sources of Power - Time Slotted Network

Total Network Power, including PPS (W)



Sources of Power – Time Slotted Network

Power Dissipation on the Processor Chip, including PPS



Summary

- Photonic time slotted networks can reduce power consumption by > order of magnitude
  - Total network power
  - Processor chip dissipation
- Further work required to understand how these networks scale under realistic data centre traffic patterns
- Transceiver power consumption is significant proportion of on-chip disipation
  - Yury will tell you more in next talk