

Scalability Analysis of optical Beneš networks based on Thermally/Electrically Tuned Mach-Zender Interferometers

Markos Kynigos, Jose Pascual, Javier Navaridas

Introduction: Problem

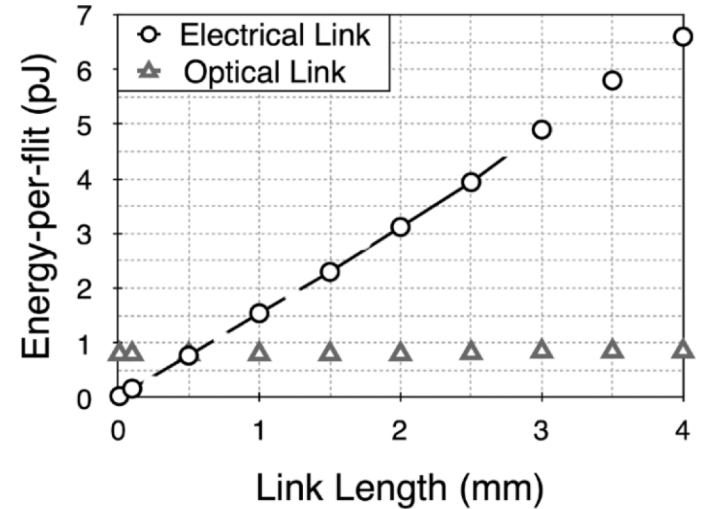


The University of Manchester

- Exa-scale & HPC: Numerous system scalability challenges
- Interconnection networks are a bottleneck for future HPC
- Scaling electrical interconnects becomes challenging:
 - Power consumption
 - Area

Introduction: Potential Solution

- Optical Interconnects based on Silicon Photonics (SiPh)
- Benefits:
 - CMOS-compatible fabrication processes
 - High bandwidth data transmission (DWDM)
 - Relatively distance-independent energy consumption



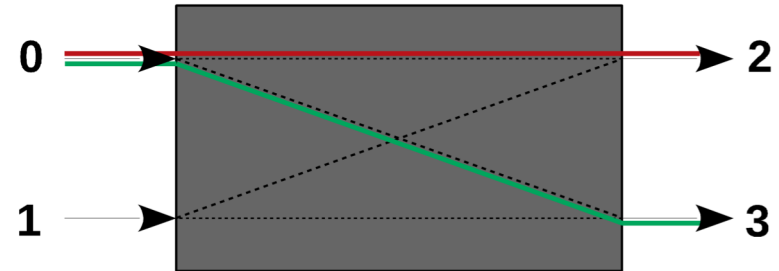
Energy difference per 64-bit flit [1]

SiPh Switch: Beneš Network & Routing

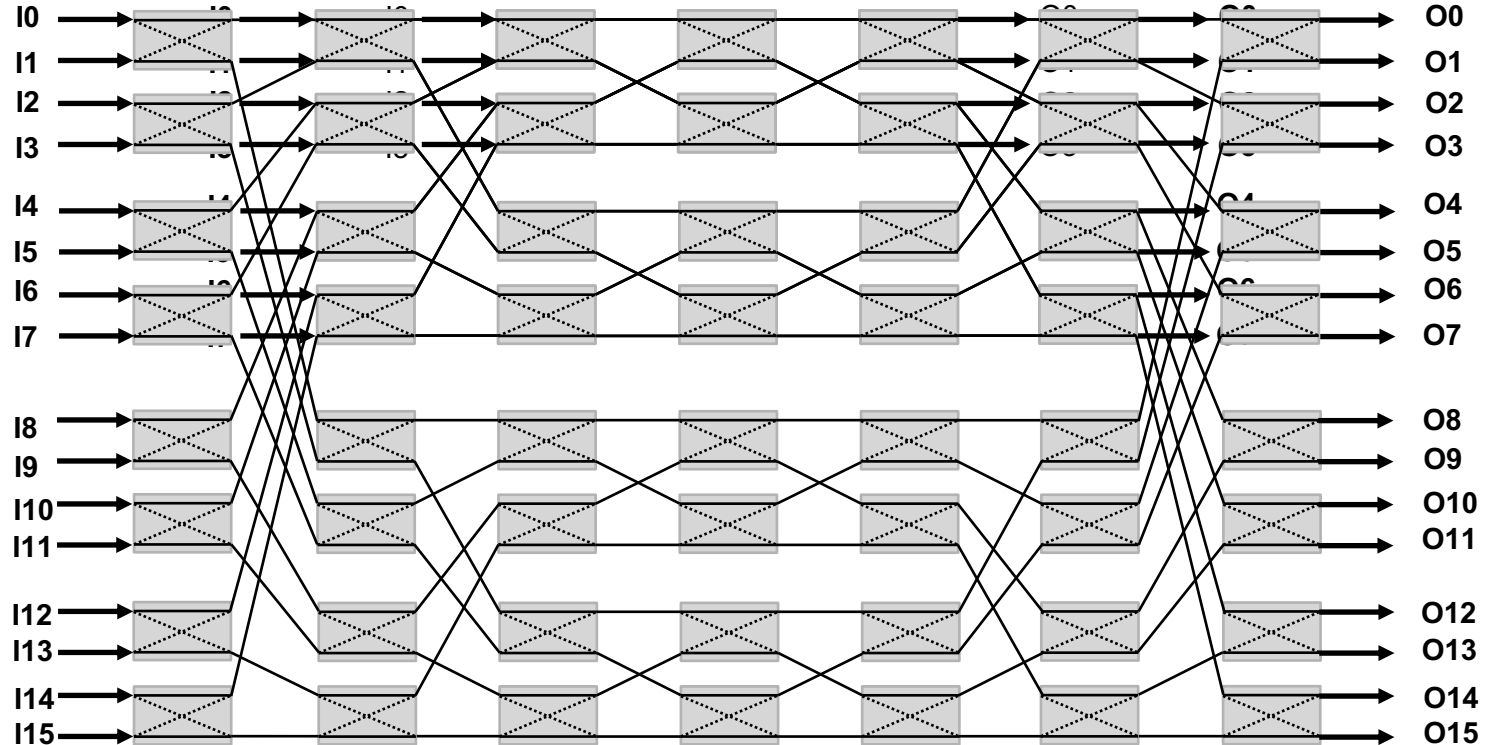
- Clos-network variant composed of 2x2 switches.
- Rearrangeably non-blocking network, minimum number of cross-points.
- Buffer-less optical communications.
- Packet switching in SiPh problematic: electro-optic conversions.
- Circuit switching with Beneš.

SiPh Switch: MZI

- Initial experimental demonstrator of 16x16 E/O tuned switch [2].
- 2x2 switch composed of 2x2 MMIs, micro-heaters and a p-i-n node.
- Switching behaviour through Mach-Zehnder Interferometry.
- Exhibits two states, “cross” and “bar”

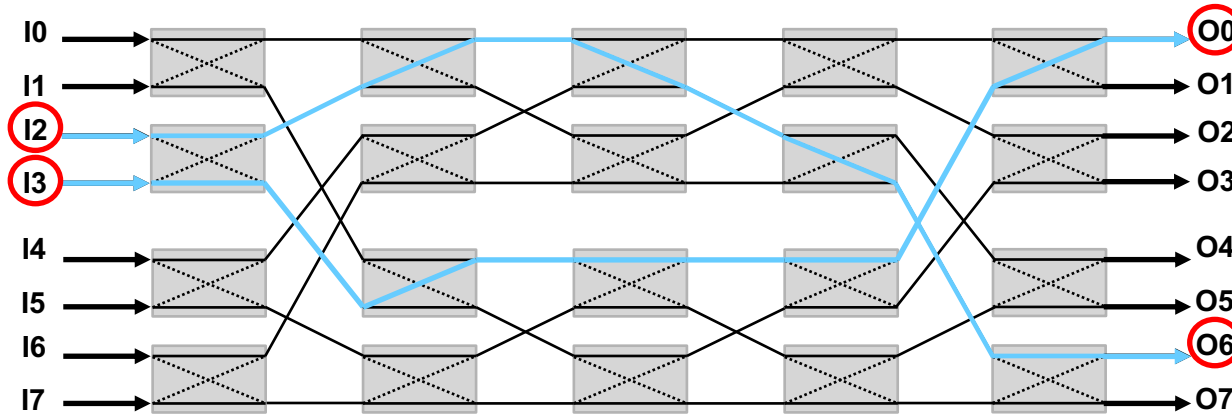


SiPh Switch



SiPh Switch: MZI Usage

- Micro-heater thermal tuning: alleviates variance due to fabrication.
- Electrical tuning: forces a state change.

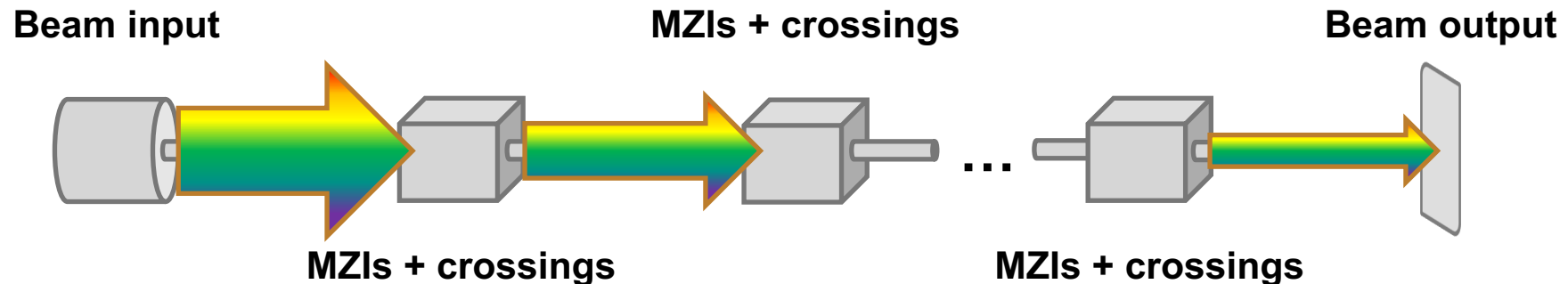


Scalability Challenges

- Goal: assessing scale-out of this network by adding more Beneš stages.
- Optical losses.
- Bit-switching energy consumption.

Scalability Challenges: Optical Loss

- Maximum Insertion Loss (max. lloss) can indicate laser specifications.
- Chief contributors: waveguides, waveguide crossings, MZIs.
- State dependency: “bar” MZIs exhibit more lloss than “cross MZIs.”



Scalability Challenges: Energy Usage

- Average energy consumption per transmitted bit (switching only).
- Contributors: MZI thermal/electrical tuning elements.

Scalability Challenges: Routing Strategies



The University of Manchester

- Devise algorithmic solutions for minimizing hardware-induced losses.
- Reduce loss/energy consumption through:
 - Loss-aware path selection.
 - State-aware path selection.
- Promising approach found in [3] for MRR-based switches.

Scalability Challenges: Routing Strategies



The University of Manchester

- **Minimise waveguide crossings:** Select paths that incur the minimum amount of crossings. Reduce max. lloss.
- **Minimise state changes:** Select paths that require the least changes in total network state. Promote MZI reuse
- **Maximise “cross” states:** Select paths that require MZIs to be in cross state. Promote MZI reuse and reduce max. lloss.
- **Baseline:** Selects a path at random after assessing path legality.

Experiment Setup: phINRFlow

- Flow-level simulator for photonic interconnects.
- Light footprint, can simulate very large scale networks.
- Support for various workloads/kernels, application traces.
 - Randomapp
 - Bisection
 - Hotregion
 - Stencil workloads based on real applications (e.g. 2-3D Meshes, Tori etc.)

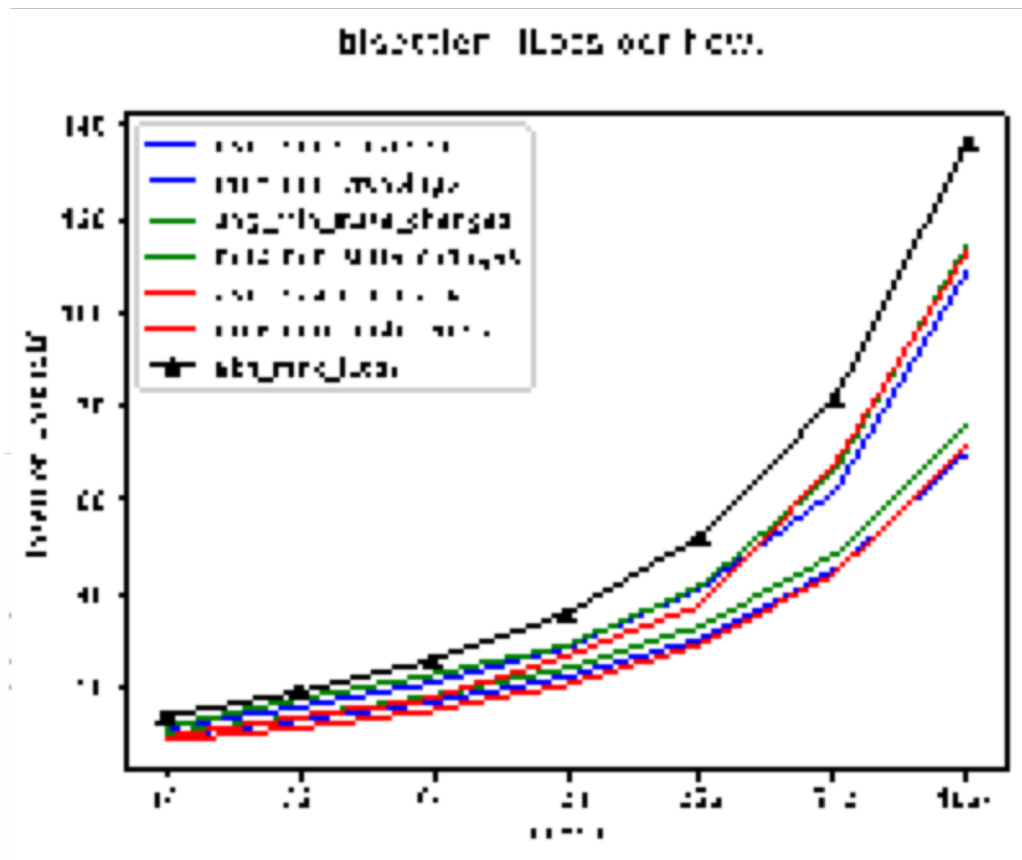
Experiment Setup: Metrics of Interest

- Maximum insertion loss (dB)
- Average insertion loss (dB)
- Bit-switching energy consumption (fJ/bit)

Component	Insertion Loss	Tuning Type	Power Cons.
Waveguide	1.18 dB/cm	Thermal	0-26 mW
Beneš Stage	0.4386 dB	Mean, STD	15.725, 6.608
Waveguide Crossing	0.05 dB	Electrical Mean, STD	3.28-5.88 mW 5.166, 0.428
“Cross” MZI	0.4 dB		
“Bar” MZI	1.4 dB		

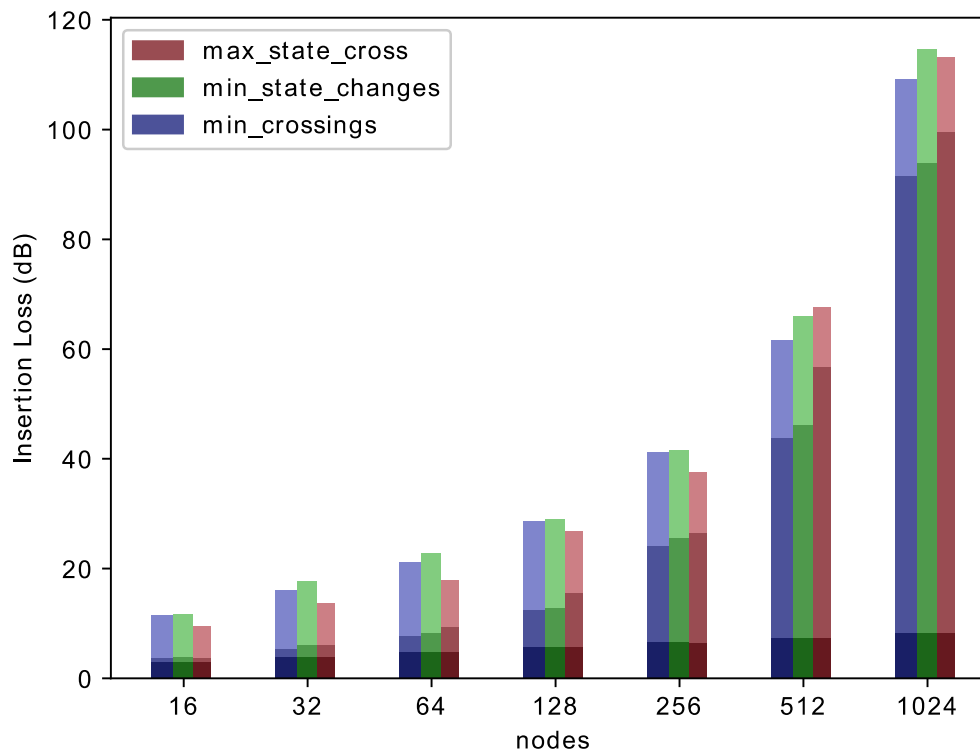
Table: Iloss & Power consumption

Evaluation: Maximum Insertion Loss



Evaluation: Insertion Loss Contributors

bisection: Max ILoss per flow.



Evaluation: Insertion Loss Contributors

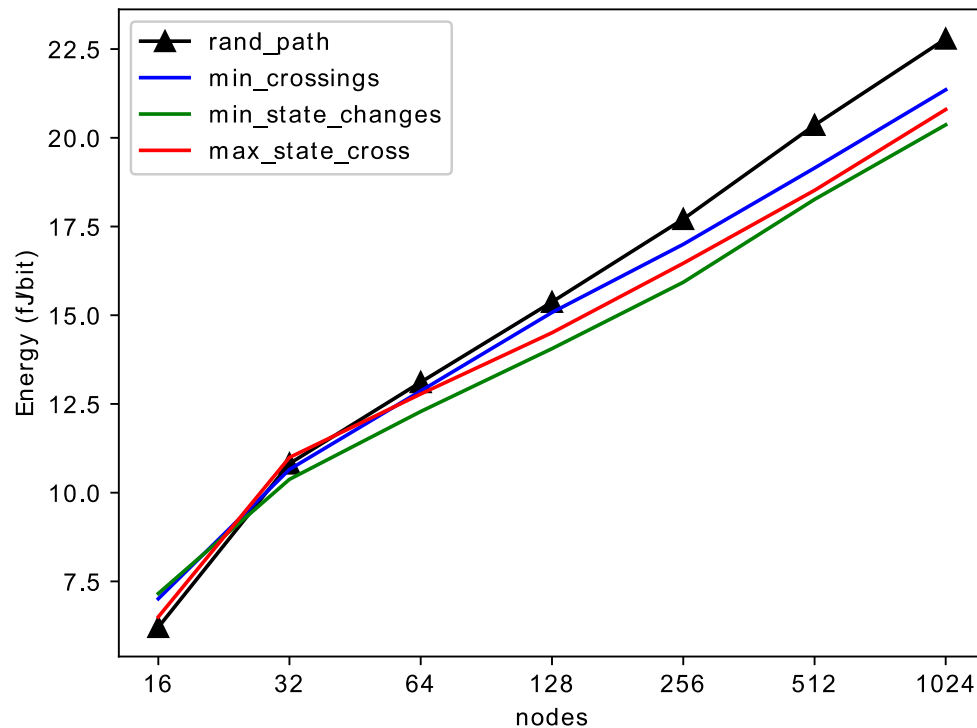


The University of Manchester

- Hybrid routing strategies
 - Minimising number of crossings and preferring MZIs in “cross” state independent criteria.
 - Can allow systems to scale up w.r.t. optical loss without adding more energy consumption.

Evaluation: Energy Consumption

bisection: Bit Switching Energy Consumption



Conclusions: Insertion Loss

- Increases in proportion to network size.
- Routing strategies reduce max. lloss., **maximising “cross” states** is best (31% reduction, best-case).
- Up to 128 endpoints: MZIs incur most lloss.
- 256-1024 endpoints: Waveguide crossings aggregate lloss
 - Primary scaling constraint: Number of crossings per path, proportion to number of endpoints, not stages.
 - Can be alleviated through chip floor-planning.

Conclusions: Energy consumption

- **Minimising waveguide crossings** is minimally worse than the baseline only under hotregion.
- **Minimising state changes** is rarely beneficial.
- **Maximising “cross” states** is the most beneficial strategy (16% reduction, best case).

Future Work

- Hybridise routing strategies
- Investigate nested topologies
 - Beneš within Beneš
 - Benes within Dragonfly

Questions.