A High-Performance Low-Power Nanophotonic On-Chip Network
Zheng Li†, Jie Wu†, Li Shang*, Alan Mickelson*, Manish Vachharajani*, Dejan Filipovic*, Wounjhang Park* and Yihe Sun†
* University of Colorado at Boulder, U.S.A.
† Tsinghua University, Beijing, China
Contact: Zheng.Li@colorado.edu

Abstract

On-chip traffic is heterogeneous, and hard to support

- Global electrical wire does not scale in terms of latency and power
- Extra power and latency overhead due to buffering, arbitration, etc
- Multicast is implemented as multiple unicasts
- 1% traffic causes 37.5% throughput degradation [Jerger et al. ISCA08]
- Difficult to have one network optimized for both traffic types

Why current NoC designs are inefficient?

- Aggressively optimized electrical network (ELE)
- Linear waveguide only hybrid photonic-electrical network (PPlanar)
- Power efficiency and performance improved by 56.1% and 63.3% over the electrical alternative, and 72.6% and 52.6% over the linear waveguide-based photonic alternative

Iris: A hybrid nanophotonic network

A power and performance efficient way to support heterogenous traffic:

- Power optimizing guidelines:
  - Minimize E/O, O/E conversions interface power
  - Minimize the number of rings and cross

Experimental results

- Compare our design (iris) to:
  - Aggressively optimized electrical network (ELE)
  - Linear waveguide only hybrid photonic-electrical network (PPlanar)
- Power efficiency and performance improved by 56.1% and 63.3% over the electrical alternative, and 72.6% and 52.6% over the linear waveguide-based photonic alternative

Conclusions

- Emerging many-core on-chip systems call for power-efficient, high-performance on-chip communication solutions
- A nanophotonic on-chip network consists of:
  - a low-latency planar-waveguide-based WDM broadcast network
  - a throughput-optimized linear-waveguide circuit-switch network
- It provides power-efficient support for both latency-critical and throughput-critical on-chip communication traffic of many-core systems
- It investigates the potentials and limitations of nanophotonic interconnect technologies

Electrical power

- Propagation loss
- S 光 loss
- Insertion time delay
- Power in linear network for multi-phase transactions required by directory protocol

Fabrication results

- Emerging many-core on-chip systems call for power-efficient, high-performance on-chip communication solutions
- A nanophotonic on-chip network consists of:
  - a low-latency planar-waveguide-based WDM broadcast network
  - a throughput-optimized linear-waveguide circuit-switch network
- It provides power-efficient support for both latency-critical and throughput-critical on-chip communication traffic of many-core systems
- It investigates the potentials and limitations of nanophotonic interconnect technologies

Conclusions

- Emerging many-core on-chip systems call for power-efficient, high-performance on-chip communication solutions
- A nanophotonic on-chip network consists of:
  - a low-latency planar-waveguide-based WDM broadcast network
  - a throughput-optimized linear-waveguide circuit-switch network
- It provides power-efficient support for both latency-critical and throughput-critical on-chip communication traffic of many-core systems
- It investigates the potentials and limitations of nanophotonic interconnect technologies

Conclusions

- Emerging many-core on-chip systems call for power-efficient, high-performance on-chip communication solutions
- A nanophotonic on-chip network consists of:
  - a low-latency planar-waveguide-based WDM broadcast network
  - a throughput-optimized linear-waveguide circuit-switch network
- It provides power-efficient support for both latency-critical and throughput-critical on-chip communication traffic of many-core systems
- It investigates the potentials and limitations of nanophotonic interconnect technologies

Conclusions

- Emerging many-core on-chip systems call for power-efficient, high-performance on-chip communication solutions
- A nanophotonic on-chip network consists of:
  - a low-latency planar-waveguide-based WDM broadcast network
  - a throughput-optimized linear-waveguide circuit-switch network
- It provides power-efficient support for both latency-critical and throughput-critical on-chip communication traffic of many-core systems
- It investigates the potentials and limitations of nanophotonic interconnect technologies