Amon: Advanced Mesh-Like Optical NoC

Sebastian Werner,
Javier Navaridas and Mikel Luján
Advanced Processor Technologies Group
School of Computer Science
The University of Manchester
Bottleneck: On-chip Interconnects in Many-core Systems

Metal Wires

• Increasing **Signal Delay** with technology scaling while gate delays decrease
• Increasing **Power Consumption** in global core-to-core interconnects due to repeaters, regenerators, or buffers
Bottleneck: On-chip Interconnects in Many-core Systems

Metal Wires

• Increasing Signal Delay with technology scaling while gate delays decrease
• Increasing Power Consumption in global core-to-core interconnects due to repeaters, regenerators, or buffers

-> Performance and Power demands cannot be met by metal wires in future many-core chips

Motivation for Optical Networks-on-chip

1. Optical data transmission by using light -> **low latency**
   (signal propagation **15ps/mm**) (global metal wire: ~262ps/mm)
2. Data can be transmitted **simultaneously** on the same waveguide at different wavelengths -> **high bandwidth**
   without adding wires
3. (Almost) **Distance independent** energy consumption
Motivation for Optical Networks-on-chip

1. Optical data transmission by using light -> **low latency** (signal propagation *15ps/mm*) (global metal wire: ~262ps/mm)
2. Data can be transmitted **simultaneously** on the same waveguide at different wavelengths -> **high bandwidth** without adding wires
3. (Almost) **Distance independent** energy consumption

Huge Potential, **BUT**: Nanophotonic components may have **high power demands**

-> Novel network architectures required to enable efficient, low-power operation
Optical on-chip Data Transmission

Wavelength: $\lambda$

Diagram:
- Laser Source
- Coupler
- Waveguide

Wavelength: $\lambda_1$
Optical on-chip Data Transmission

Wavelength: $\lambda$
Microring Resonators: Ⓢ
Ring Modulator
Laser Source
Sender A
Coupler
Waveguide
Backend Circuitry
Optical on-chip Data Transmission

Wavelength: $\lambda$

Microring Resonators: $\bigcirc$

Ring Modulator

Laser Source

Ring Filter with $\lambda_1$ resonance

Coupler

Waveguide

Photodetector

Sender A

Receiver A

Backend Circuitry
Optical on-chip Data Transmission

Wavelength: $\lambda$
Microring Resonators: $\bigcirc$

Ring Modulator
Laser Source
Coupler
Waveguide
Photodetector
Ring Filter with $\lambda_1$ resonance
Sender A
Receiver A
Backend Circuitry
Optical on-chip Data Transmission

Wavelength: $\lambda$
Microring Resonators: $\bigcirc$

Ring Modulator
Laser Source
Coupler
Waveguide
Photodetector
Ring Filter with $\lambda_1$ resonance

Sender A
Sender B
Receiver A
Receiver B

Backend Circuitry
Ring Filters for Switching (1)

Ring Filter with resonance $\lambda_2$

Waveguide 1

Waveguide 2
Ring Filters for Switching (1)

Ring Filter with resonance $\lambda_2$

- Light
- $\lambda_1$
- $\lambda_2$
- Waveguide 1
- Waveguide 2
Ring Filters for Switching (1)

Ring Filter with resonance $\lambda_2$

Waveguide 1

Waveguide 2

Light

$\lambda_1$

$\lambda_2$

Drop port
Ring Filters for Switching (2)

Number of $\lambda = \text{Number Ring Filters}$
Optical Switch for 2D Mesh
Optical Switch for 2D Mesh

Detector responding to $\lambda_3$

Detector responding to $\lambda_9$
Optical Switch for 2D Mesh

Detector responding to $\lambda_3$

Detector responding to $\lambda_9$
Optical Switch for 2D Mesh

Detector responding to $\lambda_3$

Detector responding to $\lambda_9$
ONoC Design Properties

- Network design using microring resonators is based on **deterministic** routing
- **Hardwired**, pre-defined paths between each source-destination pair

  Switching equals routing algorithm

  -> ONoC design comprises Topology, Routing algorithm and Switch architecture
Contention in Optical NoCs

\[ \lambda_1 \quad \lambda_2 \quad \lambda_3 \\
\lambda_4 \quad \lambda_5 \quad \lambda_6 \\
\lambda_7 \quad \lambda_8 \quad \lambda_9 \]
Contention in Optical NoCs

Detector responding to $\lambda_6$
Contention in Optical NoCs

Detector responding to $\lambda_6$

Ejection $\lambda_6$
Contention in Optical NoCs

Detector responding to \( \lambda 6 \)

Ejection \( \lambda 6 \)
Contention in Optical NoCs

Only one Sender per Destination at a time!
Contention in Optical NoCs

Only one Sender per Destination at a time!

Underlying **Control Network** required for destination reservation -> Req / Ack message exchange
Objectives of low-power ONoC Design

Low Laser Power

• Min. path loss -> short paths -> Low diameter
• Small $\#\lambda$ for addressing -> fewer laser sources
Objectives of low-power ONoC Design

Low Laser Power
• Min. path loss -> short paths -> Low diameter
• Small #\(\lambda\) for addressing -> fewer laser sources

Low Ring Heater Power
• Small #Microrings (20\(\mu\)W/Ring)
• Small #\(\lambda\) -> Fewer Ring Filters for Switching
State-of-the-art solutions are
1. Optical Spidergon\textsuperscript{1}
2. QuT\textsuperscript{2}

• Aim low-power
• Microring resonators
• Ring-based topology

\textsuperscript{1} S. Koohi and S. Hessabi, “Scalable architecture for a contention-free optical network on-chip,”

Optical Spidergon
Optical Spidergon
Optical Spidergon

1. (λ1)
2. (λ2)
3. (λ3)
4. (λ4)
5. (λ5)
6. (λ6)
7. (λ7)
8. (λ8)
9. (λ1)
10. (λ2)
11. (λ3)
12. (λ4)
13. (λ5)
14. (λ6)
15. (λ7)
16. (λ8)
Optical Spidergon

\[ \frac{N}{2} \lambda_s \text{ in Network for addressing} \]
\[ \rightarrow \text{Reduces Laser Power} \]
Optical Spidergon

N/2 $\lambda$s in Network for addressing
-> Reduces Laser Power
Optical Spidergon

N/2 λs in Network for addressing
-> Reduces Laser Power

Different paths to prevent overwriting data!
Optical Spidergon

1 Switch Design

(N/2 -1) Ring Filters for Switching at each node
QuT
QuT

N/4 $\lambda$s in Network for addressing
QuT

N/4 $\lambda$s in Network for addressing

2 Switch Designs (Odd/Even)

- Even Switches cheap
- Odd Switches still as expensive as in Spidergon (Ring-based Topology have similar switching demands)
Spidergon/QuT

+ N/2 and N/4 number of wavelengths in network, providing different paths to avoid contention

- Long paths in ring topologies
- Large number of ring filters for switching required
Proposal: Mesh-based Topology

Advantages over ring-topologies in oNoCs:

- **Shorter paths/diameter** than ring-based networks
- In XY Routing: At most $\sqrt{N-1}$ Ring Filters in each switch (every other node in column)
Proposal: Mesh-based Topology

Advantages over ring-topologies in oNoCs:

- **Shorter paths/diameter** than ring-based networks
- In XY Routing: At most $\sqrt{N-1}$ Ring Filters in each switch (every other node in column)

Problem:
- N number of $\lambda$s in Mesh:
  - Larger Laser Power than N/4 (QuT)
Advantages over ring-topologies in oNoCs:
- **Shorter paths/diameter** than ring-based networks
- In XY Routing: At most $\sqrt{N-1}$ Ring Filters in each switch (every other node in column)

Problem:
- N number of $\lambda$s in Mesh:
  -> Larger Laser Power than N/4 (QuT)

**Solution**: Split Mesh in 4 parts
Amon
Amon
Amon
Amon
Amon: Routing
Amon: Routing
Amon: Routing
Amon: Routing
Amon: Routing
Amon: Routing
Contention-free Routing
Contestation-free Routing
Contention-free Routing
Contestion-free Routing
Contention-free Routing
Switch Architecture

Other Switches are designed accordingly
36 Node Amon
48 Node Amon

Scaling Symmetrical to X/Y Axis
### Diameter

<table>
<thead>
<tr>
<th></th>
<th>Diameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spidergon</td>
<td>( N/4 )</td>
</tr>
<tr>
<td>QuT</td>
<td>( N/4 + 1 )</td>
</tr>
<tr>
<td>Amon</td>
<td>( (3\sqrt{N}/2) - 2 )</td>
</tr>
</tbody>
</table>

![Bar chart showing Diameter comparison between Amon and QuT for 64, 144, and 256 nodes](chart.png)
Diameter

<table>
<thead>
<tr>
<th>Diameter</th>
<th>Spidergon</th>
<th>QuT</th>
<th>Amon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N/4$</td>
<td>$N/4 + 1$</td>
<td>$(3\sqrt{N/2}) - 2$</td>
</tr>
</tbody>
</table>

Much smaller diameter with better scalability
-> shorter paths
-> less laser power
Design Configuration

- **Aim:** Low-power design, parameters are accordingly:
  - **22nm** low-voltage technology library
  - Core data rate: **4Ghz**
  - Modulator/Detector: **8Gb/s**
  - Flit Size: **16bit**
  - Standard Laser type: Laser is always on
  - Tile-width: **1mm**
  - Injection rate 0.5
  - Data is modulated on **8 wavelengths** per sender
  - Control network: Multi-Write-Single-Read Bus
  - Implementation with **DSENT**\(^1\) network modeling tool
  - 64-, 144- and 256-Node networks to assess scalability

Number of Microrings

Microrings: Modulators, Detectors, Filters

#Microrings

Savings

64 Nodes

Spideron  QuT  Amon

+ 54%  + 33%
Number of Microrings

Microrings: Modulators, Detectors, Filters

- **Savings**
  - **64 Nodes**: +54%
  - **144 Nodes**: +52%
  - **256 Nodes**: +50%

---

- **Spidergon**
  - 64 Nodes: 70,000
  - 144 Nodes: 280,000
  - 256 Nodes: 1,100,000

- **QuT**
  - 64 Nodes: 52,500
  - 144 Nodes: 210,000
  - 256 Nodes: 825,000

- **Amon**
  - 64 Nodes: 37,500
  - 144 Nodes: 140,000
  - 256 Nodes: 550,000
Number of Microrings

Microrings: Modulators, Detectors, Filters

Up to 54% savings in microrings!
Area Results

- Microring Area
- Waveguide Area

31% Savings

1.0

Spideron, QuT, Amon

64 Nodes
Area Results

- **31% Savings** for 64 Nodes
- **30% Savings** for 144 Nodes
- **29% Savings** for 256 Nodes

Bar charts showing Microring Area and Waveguide Area with the respective percentage savings for different node counts.
Power Consumption

64 Nodes
Power Consumption

- Laser Power
- Ring Heater Power
- Dynamic Power

64 Nodes

52% Savings

39%
Power Consumption

- **64 Nodes**
  - Laser Power: 39%
  - Ring Heater Power: 70%
  - Dynamic Power: 52%
  - Savings: 70%

- **144 Nodes**
  - Laser Power: 60%
  - Ring Heater Power: 70%
  - Dynamic Power: 39%
  - Savings: 60%

- **256 Nodes**
  - Laser Power: 71%
  - Ring Heater Power: 78%
  - Dynamic Power: 78%
  - Savings: 71%
Summary

Amon is a novel mesh-based optical NoC comprising topology, switch architecture and routing algorithm
Summary

Amon is a novel mesh-based optical NoC comprising topology, switch architecture and routing algorithm.

Compared to ring-based Spidergon and QuT, Amon saves:

- Laser Power:
  - Short paths -> lower path losses
  - N/4 Wavelengths in Network
- Ring Heater Power:
  - Fewer Ring filters for switching -> less ring tuning required
- Total Power Savings up to 78% / 71%
- Area due to fewer microrings (up to 31% / 18%)
- Mesh Structure suitable for tile-based VLSI implementation
Thank you!
Questions?
Zero Load Latency

Control Network:
• Packet Size 2bit for packet type (req/ack/nack)
• 4Ghz Core clk and 8Gb/s Modulator: 2 bits per clock clk
• Total latency: Modulation (1 cycle) + On-the-fly (1 cycle) + Detection (1 cycle)
  = 3 cycles
• Destination checking: 6 cycles (req + ack)
Zero Load Latency

Control Network:
- Packet Size 2bit for packet type (req/ack/nack)
- 4Ghz Core clk and 8Gb/s Modulator: 2 bits per clock clk
- Total latency: Modulation (1 cycle) + On-the-fly (1 cycle) + Detection (1 cycle) = 3 cycles
- Destination checking: 6 cycles (req + ack)

Data Network:
- Assuming 128bit data packet
- Data transmission with 8 modulators: 128 / 8 / 2 = 8 cycles for modulation, 1 on-the-fly, 8 for detection -> 17 cycles
- Total: 23 Cycles
Zero Load Latency

**Control Network:**
- Packet Size 2bit for packet type (req/ack/nack)
- 4Ghz Core clk and 8Gb/s Modulator: 2 bits per clock clk
- Total latency: Modulation (1 cycle) + On-the-fly (1 cycle) + Detection (1 cycle) = 3 cycles
- **Destination checking:** 6 cycles (req + ack)

**Data Network:**
- Assuming 128bit data packet
- Data transmission with 8 modulators: 128 / 8 / 2 = 8 cycles for modulation, 1 on-the-fly, 8 for detection -> **17 cycles**
- **Total:** 23 Cycles

- with 200ps clock cycle and 15ps/mm propagation delay, every destination within 18 hops is reached in one clock cycle
- Larger network size has insignificant impact on latency
- Adding modulators or using faster ones (up to 40Gb have been fabricated) further decreases latency
# Insertion Loss Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laser efficiency</td>
<td>5 dB</td>
</tr>
<tr>
<td>Coupler loss</td>
<td>1 dB</td>
</tr>
<tr>
<td>Waveguide propagation loss</td>
<td>100 dB/m</td>
</tr>
<tr>
<td>Ring: Through loss</td>
<td>0.01 dB</td>
</tr>
<tr>
<td>Ring: Drop loss</td>
<td>1 dB</td>
</tr>
<tr>
<td>Modulator Insertion Loss</td>
<td>1 dB</td>
</tr>
<tr>
<td>Modulator Extinction</td>
<td>1 dB</td>
</tr>
<tr>
<td>Photodetector loss</td>
<td>1 dB</td>
</tr>
</tbody>
</table>
Control Network MWSR

**Power:**
21%, 19%, and 17% of Amon (64, 144, 256 Nodes)
Only 1 Modulator compared to 8 leads to small ring heater power and area

**Waveguide Area** becomes significant as one waveguide reaching to every other node in the oNoC is added for each node
Control Network
Control Network

- Req - Ack/NegAck messages for destination reservation
Control Network

- **Req - Ack/NegAck** messages for destination reservation
- Commonly implemented as a **Multiple-Write-Single-Read bus**
Technology Parameters Area

Waveguide->Pitch = 4e-6 # m
Ring->Area = 100e-12 # m2
Photodetector->Area = 10e-12 # m2
Power Consumption

Amon total power:

64 Nodes: 0.83W
144 Nodes: 4W
256 Nodes: 15W
Area Results
Area Results

- **Microring Area**
- **Waveguide Area**

```
<table>
<thead>
<tr>
<th>Node</th>
<th>Microring Area</th>
<th>Waveguide Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spideron</td>
<td>3.5</td>
<td>1.75</td>
</tr>
<tr>
<td>QuT</td>
<td>3.5</td>
<td>1.75</td>
</tr>
<tr>
<td>Amon</td>
<td>1.75</td>
<td>3.5</td>
</tr>
</tbody>
</table>
```

64 Nodes
Area Results

![Graph showing area results for Microring Area and Waveguide Area for 64 and 144 Nodes.](image-url)
Area Results

Microring Area  Waveguide Area

mm$^2$

<table>
<thead>
<tr>
<th></th>
<th>64 Nodes</th>
<th>144 Nodes</th>
<th>256 Nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spidergon</td>
<td>7</td>
<td>20</td>
<td>37.5</td>
</tr>
<tr>
<td>QuT</td>
<td>3.5</td>
<td>9.5</td>
<td>12.5</td>
</tr>
<tr>
<td>Amon</td>
<td>1.75</td>
<td>5.25</td>
<td>25</td>
</tr>
</tbody>
</table>

64 Nodes

144 Nodes

256 Nodes
Power Consumption

- Laser Power
- Ring Heater Power
- Dynamic Power

WATTS

Spidergon | QuT | Amon
---|---|---
0.5 | 1.0 | 0.5

64 Nodes
Power Consumption

- Laser Power
- Ring Heater Power
- Dynamic Power

WATTS

64 Nodes
- Spidergon: 1.5 W
- QuT: 0.5 W
- Amon: 0.5 W

144 Nodes
- Spidergon: 10.5 W
- QuT: 3.5 W
- Amon: 1 W
Power Consumption

- Laser Power
- Ring Heater Power
- Dynamic Power

**64 Nodes**
- Spidergon: 1.5 W
- QuT: 1 W
- Amon: 0.5 W

**144 Nodes**
- Spidergon: 10.5 W
- QuT: 3.5 W
- Amon: 7 W

**256 Nodes**
- Spidergon: 52.5 W
- QuT: 35 W
- Amon: 17.5 W
VLSI Layout: Shared Laser Sources

Laser Sources

Coupler

Splitter

λ1

λ2

λn

WG1

WG2

WG3

WG4
VLSI Layout: Shared Laser Sources

- Bypass/Intermesh Link
- Mesh link
- Laser Sources + Couplers + Splitters

λ1 ... λ(N/4)
# Amon: Evaluation & Comparison

For comparison:

<table>
<thead>
<tr>
<th>#Nodes</th>
<th>Spidergon</th>
<th>QuT</th>
<th>Amon</th>
<th>CN</th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>4.4e-06</td>
<td>3.7e-06</td>
<td>2.7e-06</td>
<td>7.0e-07</td>
</tr>
<tr>
<td>144</td>
<td>1.44e-05</td>
<td>1.2e-05</td>
<td>9.2e-06</td>
<td>2.3e-06</td>
</tr>
<tr>
<td>256</td>
<td>3.48e-05</td>
<td>2.9e-05</td>
<td>2.3e-05</td>
<td>5.0e-06</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>#Nodes</th>
<th>Spidergon</th>
<th>QuT</th>
<th>Amon</th>
<th>CN</th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>1.68e-06</td>
<td>1.4e-06</td>
<td>1.5e-06</td>
<td>3.1e-06</td>
</tr>
<tr>
<td>144</td>
<td>3.84e-06</td>
<td>3.2e-06</td>
<td>3.6e-06</td>
<td>9.1e-06</td>
</tr>
<tr>
<td>256</td>
<td>6.7e-06</td>
<td>5.5e-06</td>
<td>6.6e-06</td>
<td>4.3e-05</td>
</tr>
</tbody>
</table>

### Microring area (m²)

### Waveguide area (m²)

### Total area normalized to Amon

<table>
<thead>
<tr>
<th>#Nodes</th>
<th>Amon</th>
<th>QuT</th>
<th>Spidergon</th>
<th>CN</th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>1</td>
<td>1.21</td>
<td>1.44</td>
<td>0.88</td>
</tr>
<tr>
<td>144</td>
<td>1</td>
<td>1.19</td>
<td>1.42</td>
<td>0.94</td>
</tr>
<tr>
<td>256</td>
<td>1</td>
<td>1.16</td>
<td>1.40</td>
<td>1.61</td>
</tr>
</tbody>
</table>

For comparison:
eNoC 64-node Mesh: Area: 1.77e-06 (~ 40% of Amon)
44 injection channels for destinations in
< N/4 (left/right)
> N/4 (left/right)
hop distance

N/4 wavelengths in network
-> less switching rings
-> Same #modulators at each node

But:
Ring topology causes long paths leading to high IL