Network on Chip Architectures

BY
JAGAN MURALIDHARAN
NIRAJ VASUDEVAN
Multi-Core Chips

- No more single processor systems
- High computational power requirements
- Increasing clock frequency increases power dissipation
- Power consumption increases by 60% for every 400 MHz increase in clock frequency
  - Solution: Introduce core level parallelism; distribute tasks to multiple cores
- Intel 48 core single chip cloud computer processor, Intel 80 core processor
- **Interconnection of cores**
Interconnects
Interconnects

- Most commonly used on-chip interconnect architecture is the shared bus

Cons:
- Delay becomes a problem with more number of cores
- Not scalable
Introduction to NOC

- Packet switching based communication network
- Information is routed via routers and interconnect links
- Desired Properties
  - Scalability
  - Higher bandwidth
  - Less contention
  - Shortest path between any two nodes
NOC

- The challenge is to develop a scalable infrastructure for communication
- Packet based communication between cores
- Flit based wormhole switching
Metrics

- Performance Metrics:
  - Throughput: bits successfully transferred to destination per core per second (bps per core)
  - Packet Energy: Average energy dissipation to reach destination
  - Transport latency: the time (in clock cycles) that elapses from between the occurrence of a message header injection into the network at the source node and the occurrence of a tail flit reception at the destination node
Small World Topology

- Inspired from nature – microbial colonies and social networks
- Has many short links and a few long links

\[
P_{ij} = \frac{h_{ij}}{\sum_{i,j} h_{ij}}
\]

\[
P(i, j) = \frac{l_{ij}^{-\alpha} f_{ij}}{\sum_{i=1}^{n} \sum_{j=1}^{n} l_{ij}^{-\alpha} f_{ij}}
\]
Upcoming Types of Interconnects

- Traditional wired interconnects are highly power hungry
- 3D Interconnects
- Photonic Interconnects
- Wireless Interconnects
3D Interconnects

- Much lesser interconnect lengths
- Thus improved performance and power consumption
Photonic Interconnects

- **Pros:**
  - Greater bandwidth
  - Low power

- **Cons:**
  - Fabrication
  - Scalability
Wireless Interconnects

- Uses on-chip wireless antennas
  - Carbon nanotube antennas
  - Mm-wave antennas

- Advantage:
  - Low power dissipation
  - Only a single hop is required for intra-subnet communication

- Disadvantage:
  - Noisy wireless channel
Thermal Hotspots

- Reasons for *hotspots*:
  - High traffic density
  - High power density
  - Imbalanced and localized traffic
Dynamically Reconfigurable NOC

- Traditionally, routes are pre-calculated offline
  - Dijkstra’s algorithm

- Frequently used switches will have:
  - High traffic density
  - High power
  - Increase in NOC temperature
Distance Vector Routing

- A distributed algorithm – no single authority
- Every node maintains path costs (distance vectors) to all other nodes
- When NOC temperature goes above threshold, each node recalculates its vector table

\[ C_{i,j} = \begin{cases} 1, & T_{link_{i,j}} < T_{link}^{th} \text{ and } T_{switch_{j}} < T_{switch}^{th} \\ \infty, & T_{link_{i,j}} > T_{link}^{th} \text{ or } T_{switch_{j}} > T_{switch}^{th} \end{cases} \]
Comparison of Architectures

- Packet Energy per Bandwidth (nJ/TBps) for different NoC architectures with scaling of system size

<table>
<thead>
<tr>
<th>System Size</th>
<th>Flat Mesh</th>
<th>WiNoC_{THz}</th>
<th>WiNoC_{subTHz}</th>
<th>RFNoC</th>
<th>Clos</th>
<th>Photonic NoC</th>
</tr>
</thead>
<tbody>
<tr>
<td>128</td>
<td>560.0</td>
<td>21.0</td>
<td>31.3</td>
<td>34.1</td>
<td>16.6</td>
<td>34.1</td>
</tr>
<tr>
<td>256</td>
<td>721.7</td>
<td>13.3</td>
<td>19.6</td>
<td>25.8</td>
<td>9.0</td>
<td>29.6</td>
</tr>
<tr>
<td>512</td>
<td>1171</td>
<td>10.0</td>
<td>15.4</td>
<td>32.8</td>
<td>7.3</td>
<td>35.1</td>
</tr>
</tbody>
</table>
Comparison of Interconnect paradigms

<table>
<thead>
<tr>
<th>Design Requirements</th>
<th>3D Integration</th>
<th>Optical Interconnects</th>
<th>Wireless Links</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance Gains</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bandwidth Advantage</td>
<td>Multiple layers with active devices</td>
<td>Silicon photonic components</td>
<td>On-chip metal or CNT-based antennas</td>
</tr>
<tr>
<td>Lower Power Dissipation</td>
<td>Higher connectivity &amp; less hop count</td>
<td>High speed optical devices and links</td>
<td>Direct point-to-point wireless links between smaller subnets</td>
</tr>
<tr>
<td>Reliability</td>
<td>Shorter average path length</td>
<td>Negligible power dissipation in optical data transport</td>
<td>Multi-hop paths replaced by single hop links</td>
</tr>
<tr>
<td>Challenges</td>
<td>Vertical Via Failure</td>
<td>Temperature sensitivity of photonic components</td>
<td>Noisy wireless channel</td>
</tr>
<tr>
<td></td>
<td>Heat dissipation due to higher power density, yield</td>
<td>Integration of on-chip photonic components</td>
<td>Low power mm-wave transceivers &amp; Control over CNT growth</td>
</tr>
</tbody>
</table>
Comparison of Achievable Bandwidth

- Wireless NoC outperforms other architectures
Comparison of Area Overheads

- Photonic and Clos NoC have the highest area overheads
References


[2] Luca Benini, Giovanni De Micheli. “Networks on Chips: A New SoC Paradigm”.


