CNOC: High-Radix Clos Network-on-Chip

Yu-Hsiang Kao
Advisor – H. Jonathan Chao
Outline

• Introduction
  – Chip Multiprocessors (CMP)
  – Previous Network-on-Chip (NoC) Topologies and their respective problems
  – Clos Network-on-Chip (CNOC)
• Research Issues
• Proposed solutions
  – The scheduling algorithm of CNOC
  – The design of the Hierarchical Round-Robin Arbiter (HRRA)
  – The floorplanning methodology of CNOC
• Experimental results
• Conclusion
Chip Multiprocessor (CMP)

- Billions of transistors on chip
  - CMP exploits parallelism to scale up the processor performance.
- Logical view of a shared-memory CMP
  - Each Processing Element (PE) has its own L1 cache.
  - All PEs share a logical L2 cache
- Performance Bottleneck
  - L2 cache bandwidth
  - Bus bandwidth
Some CMPs from Industry

- Intel 80 core teraflops research chip
  - 65nm, year 2007
  - [http://techresearch.intel.com/articles/Tera-Scale/1449.htm](http://techresearch.intel.com/articles/Tera-Scale/1449.htm)

- Sun ROCK, 16-core SPARC processor
  - 65nm, year 2008
Some CMPs from Industry

- IMB, Sony, Toshiba, Cell Processor
- 65nm year 2007

- Tilera TILEPro64 Processor
- 90nm, year 2008
- http://www.tilera.com/products/TILEPro64.php
NoC Topologies

- Low-radix
  - 2D Mesh
- Problems
  - Low bi-section bandwidth
  - Low throughput under uniform traffic
  - High zero-load latency
  - Low power efficiency
NoC Topologies

- Concentrated Mesh and Concentrated MeshX2
  - High-radix
  - Improved zero-load latency
  - Throughput is not improved.
NoC Topologies

- Flattened Butterfly
  - High-radix
  - Improved zero-load latency
  - Improved throughput
  - Router radix is 10.

- Area/power/frequency issues
NoC Topologies

- Fat-tree
  - High-radix
  - Improved zero-load latency
  - Improved throughput
  - Large number of routers
- Area/power issues

NoC Topologies

- Clos Network-on-Chip (CNOC)
  - High-radix
  - Improved zero-load latency
  - Improved throughput
  - Improved power efficiency
Research Issues and Contributions

• CNOC composed of the conventional VC routers can only achieve 65% throughput under uniform traffic.
  – We proposed a new scheduling algorithm for CNOC and improved the throughput under uniform traffic from 65% to 78%.
• The conventional router design cannot achieve 2GHz because of the long critical path delay of the traditional round-robin arbiter.
  – We designed a hierarchical round-robin arbiter and verified its delay and power performance after place and route using SOC Encounter.
• The long interconnects of the CNOC might cause delay and power bottlenecks.
  – We proposed a floorplanning methodology for CNOC.
Outline

• Introduction
  – Chip Multiprocessors (CMP)
  – Previous Network-on-Chip (NoC) Topologies and their respective problems
  – Clos Network-on-Chip (CNOC)

• Research Issues
• Proposed solutions
  – The scheduling algorithm of CNOC
  – The design of the Hierarchical Round-Robin Arbiter (HRRA)
  – The floorplanning methodology of CNOC

• Experimental results
• Conclusion
Wormhole Routing and Virtual Channels

• Wormhole routing
  – Different packets cannot interleave in the same input queue.

• Virtual Channel (VC) flow control
  – Different packets can interleave on the same physical link.
Problems of the Conventional Router Design

- Head-of-Line blocking problem
  - Virtual Output Queue (VOQ) buffer structure
The Conventional Scheduling Algorithm

- The input-first scheduling algorithm
  - Dual Round-Robin Matching (DRRM)
  - The round-robin arbiters update their pointers in every clock cycle.
Problems of the Conventional Router Design

• Tail-of-Line blocking problem
  – Packet-mode Dual Round Robin Matching (PDRRM) algorithm
  – The IAs/OAs keep granting a packet until its tail flit is transferred.

<table>
<thead>
<tr>
<th>Upstream Router</th>
<th>Downstream Router1</th>
<th>Downstream Router2</th>
</tr>
</thead>
<tbody>
<tr>
<td>VOQ1 VOQ2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IA: input arbiter</td>
<td>OA: output arbiter</td>
<td>OA: output arbiter</td>
</tr>
<tr>
<td>VOQ1 VOQ2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IA: input arbiter</td>
<td>OA: output arbiter</td>
<td>OA: output arbiter</td>
</tr>
</tbody>
</table>
The Performance of PDRRM

- In-house cycle-accurate simulator
- 8 VCs/VOQs in each input port
- Each input port can accommodate 256 flits
- 8-flit packet size
- Throughput under uniform traffic
  - 65% -> 81%
Outline

• Introduction
  – Chip Multiprocessors (CMP)
  – Previous Network-on-Chip (NoC) Topologies and their respective problems
  – Clos Network-on-Chip (CNOC)
• Research Issues
• Proposed solutions
  – The scheduling algorithm of CNOC
  – The design of the Hierarchical Round-Robin Arbiter (HRRA)
  – The floorplanning methodology of CNOC
• Experimental results
• Conclusion
The Micro-architecture of the PDRRM Router
Generic RRA and Hierarchical RRA

• An 8x8 generic RRA has a 447ps critical path delay.
  – Too close to 500ps
  – Hard to scale up the operating frequency
The Performance of HRRA

- The VHDL code of both designs are synthesized and analyzed by the Cadence Encounter RTL Compiler using STMicroelectronics Company 65nm technology.
- SOC Encounter is used to place and route these designs.
- Power evaluation is performed based on an activity factor of 100% using Cadence Encounter.

<table>
<thead>
<tr>
<th></th>
<th>Radix</th>
<th>Critical Path Delay (ps)</th>
<th>Power (mW per GHz)</th>
<th>Area (um²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generic RRA</td>
<td>4</td>
<td>241</td>
<td>0.26</td>
<td>282</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>477</td>
<td>0.704</td>
<td>698</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>637</td>
<td>1.73</td>
<td>1586</td>
</tr>
<tr>
<td>HRRA</td>
<td>8</td>
<td>369 (-22.9%)</td>
<td>0.8 (+13.6%)</td>
<td>726 (+4%)</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>487 (-23.5%)</td>
<td>1.44 (-16.8)</td>
<td>1146 (-27.7%)</td>
</tr>
</tbody>
</table>

Floorplan and layout of a radix-8 VOQ PDRRM router (SRAMs not included).
Outline

• Introduction
  – Chip Multiprocessors (CMP)
  – Previous Network-on-Chip (NoC) Topologies and their respective problems
  – Clos Network-on-Chip (CNOC)
• Research Issues
• Proposed solutions
  – The scheduling algorithm of CNOC
  – The design of the Hierarchical Round-Robin Arbiter (HRRA)
  – The floorplanning methodology of CNOC
• Experimental results
• Conclusion
Floorplanning

- Routing over resources or routing on dedicated channels

- Wiring density
  - Wiring density is the maximum number of tile-to-tile wires routable across a tile edge.
  - 40-60% of the length of a tile edge is routable for global interconnects.
Problem Definition

- \( NxN \) identical tiles \( T \)
- Routers \( R = \{ R_i, i = 1,2,3 \ldots n \} \)
- Links \( L = \{ L_i, i = 1,2,3 \ldots m \} \)
- \( M(L_i) \) as the Manhattan distance between the two nodes that \( L_i \) connects
- \( E_k \) is the energy that a flit consumes when it travels through a link with Manhattan distance \( k \)
- Objective function:

\[
\Phi(R, L) = \sum_{i=1}^{m} E_{M(L_i)}
\]
Floorplan Result

- Under the constraints:
  - Longest wire length is 7 hops
  - 40% length of the tile edge is routable
  - Any link can contain up to 128 bits
  - All the wires can be operated under 2 GHz
  - There could be multiple solutions
Outline

• Introduction
  – Chip Multiprocessors (CMP)
  – Previous Network-on-Chip (NoC) Topologies and their respective problems
  – Clos Network-on-Chip (CNOC)
• Research Issues
• Proposed solutions
  – The scheduling algorithm of CNOC
  – The design of the Hierarchical Round-Robin Arbiter (HRRA)
  – The floorplanning methodology of CNOC
• Experimental results
• Conclusion
Simulation Setup

<table>
<thead>
<tr>
<th>Router Design</th>
<th>2-D Mesh</th>
<th>CMeshX2</th>
<th>Fat-tree</th>
<th>CNOC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DRRM VC</td>
<td>DRRM VC</td>
<td>DRRM VC</td>
<td>PDRRM VOQ</td>
</tr>
<tr>
<td>Total router Number</td>
<td>36 radix-5, 24 radix-4, 4 radix-3</td>
<td>32 radix-8</td>
<td>32 radix-8, 16 radix-4</td>
<td>24 radix-8</td>
</tr>
<tr>
<td>Pipeline stage #</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>VC number</td>
<td>4</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Frequency</td>
<td>2GHz</td>
<td>2GHz</td>
<td>2GHz</td>
<td>2GHz</td>
</tr>
<tr>
<td>Input port memory size</td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td>Flit size (bit)</td>
<td>32</td>
<td>32</td>
<td>32</td>
<td>32</td>
</tr>
</tbody>
</table>

*For energy consumption on NOC links, we use Cadence Spectre to run simulations on models with parameters given by the Predictive Technology Model (PTM) of various lengths. The energy consumed by a flit is 2.16 pJ per 2mm.

50ns Throughputs

- 50-ns throughput of different topologies under different kinds of synthetic traffic patterns.
Low-load Packet Latencies

- Average packet latencies under 5% injection rate of different kinds of synthetic traffic patterns.
Normalized Power Efficiencies

- **Power Efficiency** = \(1/E\), where \(E = (\text{time for each PE to finish sending 1000 packets at the injection rate with an average latency of 50ns}) \times (\text{total energy dissipated during the process})\).
Low-load Power Consumption

- Power consumption of different topologies under 5% traffic load.
Area Comparison

- The NOC area percentage over the 16mm by 16mm chip area
Benchmark Simulation Setup

- We use Graphite, an x86 multicore simulator, to generate the SPLASH-2 traffic traces.
- Then we injected the traffic traces into our in-house cycle-accurate simulator to obtain the delay and power performance.
- In Graphite, it is assumed that there are 64 PEs, each with its own private L1/L2 (32kB/512kB) caches.
- The caches are coherent with MSI protocol, and the packet sizes are either 8-byte or 64-byte (2 flit or 16 flits)

<table>
<thead>
<tr>
<th>Benchmark</th>
<th>Experimental data set</th>
<th>Default data set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barnes</td>
<td>64k</td>
<td>16k</td>
</tr>
<tr>
<td>Cholesky</td>
<td>tk29.0</td>
<td>tk15.0</td>
</tr>
<tr>
<td>FFT</td>
<td>16M</td>
<td>64k</td>
</tr>
<tr>
<td>FMM</td>
<td>64k</td>
<td>16k</td>
</tr>
<tr>
<td>LU</td>
<td>2048x2048</td>
<td>512x512</td>
</tr>
<tr>
<td>Ocean</td>
<td>2050x2050</td>
<td>258x258</td>
</tr>
<tr>
<td>Radiosity</td>
<td>roomlarge</td>
<td>Room</td>
</tr>
<tr>
<td>Radix</td>
<td>1M</td>
<td>1M</td>
</tr>
<tr>
<td>Raytrace</td>
<td>balls4</td>
<td>Car</td>
</tr>
<tr>
<td>Volrend</td>
<td>head</td>
<td>head</td>
</tr>
<tr>
<td>Water-SP</td>
<td>32k</td>
<td>512</td>
</tr>
</tbody>
</table>
Benchmark Simulation Results

- Average packet latencies of the SPLASH-2 Benchmarks
Conclusion

• We proposed a new scheduling algorithm for CNOC and improved the throughput with the conventional scheduling algorithm under uniform traffic from 65% to 78%.

• We designed a hierarchical round-robin arbiter and verified its delay and power performance after place and route using SOC Encounter.

• We proposed a floorplanning methodology for CNOC.

• We compared CNOC with 2D Mesh, CMeshX2, and Fat-tree in terms of 50-ns throughput, low-load latency, power efficiency, low-load power consumption, and average packet latencies under SPLASH-2 benchmarks.

• The works have been published in
  – IEEE/ACM NOCS 2010
Thank You!