ARIADNE: Agnostic Reconfiguration In A Disconnected Network Environment

Konstantinos Aisopos, Andrew DeOrio, Li-Shiuian Peh, and Valeria Bertacco
2011 PACT

Presented by: Helen Arefayne Hagos, Yajing Chen
Motivation

- Permanent failure

- Time

- 150 M transistors
- 800 M transistors
- 1B transistors
Solution

Reconfiguration
## Existing solutions

<table>
<thead>
<tr>
<th>Targeted fault</th>
<th>Reliability</th>
<th>Performance</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bounded in number</td>
<td>Limited</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Pattern constrained</td>
<td>Limited</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Not constrained Off-chip</td>
<td>High</td>
<td>Good</td>
<td>Excess</td>
</tr>
<tr>
<td>On-chip</td>
<td>High</td>
<td>Bad</td>
<td>High</td>
</tr>
<tr>
<td>Immunet</td>
<td>High</td>
<td>Bad</td>
<td>Low</td>
</tr>
<tr>
<td>Vicis</td>
<td>Limited</td>
<td>Bad</td>
<td>Low</td>
</tr>
</tbody>
</table>
Proposed solution

- Agnostic
- Reconfiguration algorithm
- In
- A
- Disconnected
- Network
- Environment
1ST Broadcast

Conflict Resolution:
2\textsuperscript{nd} Broad Cast...

- \text{RT}[2]=E
- \text{RT}[2]=E
- \text{RT}[2]=W

- \text{up}
- \text{up}
- \text{down}
- \text{down}

- \text{RT}[2]=N, \text{ERT}[2]=N
- \text{RT}[2]=N

- N cycles for a single broadcast
- O(N^2) cycles for the entire reconfiguration
5th Broadcast


RT[5]=S


Illegal path from Node 0 to Node 5
SYNCHRONIZATION

- Atomic broadcasts

- Atomic synchronization
## Experiment Setup

- ARIADNE is implemented in the Wisconsin Multifacet GEMS simulator as part of the GARNET network model
- Two state-of-art routing algorithms: Vicis and Immunet

<table>
<thead>
<tr>
<th>Network Architecture</th>
<th>System Configuration</th>
<th>Simulation Input</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GARNET</strong></td>
<td><strong>GEMS</strong></td>
<td></td>
</tr>
<tr>
<td>Network Topology</td>
<td>8x8 2D mesh</td>
<td></td>
</tr>
<tr>
<td>Memory Controllers</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Channel Width</td>
<td>64 bits</td>
<td></td>
</tr>
<tr>
<td>Router Architecture</td>
<td>5-stage pipeline</td>
<td></td>
</tr>
<tr>
<td>Router Ports, VCs</td>
<td>5, 2(private)</td>
<td></td>
</tr>
<tr>
<td>Router Buffers/port</td>
<td>5-flit/VC</td>
<td></td>
</tr>
<tr>
<td>Processors</td>
<td>in-order SPARC cores</td>
<td></td>
</tr>
<tr>
<td>Coherence</td>
<td>MOESI protocol</td>
<td></td>
</tr>
<tr>
<td>L1 Caching</td>
<td>private unified 32KB/node 2-way latency:3cycles</td>
<td></td>
</tr>
<tr>
<td>L2 Caching</td>
<td>shared distributed 1MB/node 16-way latency:15cycles</td>
<td></td>
</tr>
<tr>
<td>Synthesis Traffic</td>
<td></td>
<td>Uniform Random(UR) TRanspose (TR)</td>
</tr>
<tr>
<td>Benchmark Suite</td>
<td></td>
<td>PARSEC</td>
</tr>
<tr>
<td>Packet Length</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Simulation Time</td>
<td></td>
<td>100K cycles</td>
</tr>
<tr>
<td>Simulation Warmup</td>
<td></td>
<td>10K cycles</td>
</tr>
</tbody>
</table>

University of Michigan
Performance Evaluation: Synthetic Traffic

- **Average Latency**: the delay experienced by a packet from source to destination
- **Throughput**: the rate of packets delivered per cycle
- **Zero Load Latency**

![Graphs showing Average Latency and Saturation Throughput over Injected Faults](image-url)
Performance Evaluation: PARSEC

- Average Latency for 50 faults

- Latency over variable number of faults
Reliability Evaluation

- The ability to maximize the connectivity of a faulty network
- Synthetic Traffic
- PARSEC
- The probability of deadlock
Reconfiguration Evaluation

- During reconfiguration, all packets experience an additional delay of $N^2$ cycles (N=64; 4096 cycles)
- Average Latency over time

---

**Average Latency**

- **0 to 1 faults**
- **0 to 10 faults**
- **0 to 50 faults**
- Reconfiguration initiated
- Reconfiguration completed

---

**Reconfiguration Duration**

- zero load (0.01 flits/node/cycle)
- 50% capacity (0.05 flits/node/cycle)
- 100% capacity (0.1 flits/node/cycle)
# Hardware Overhead

- Verilog HDL
- Beginning with 5-stage router design: 2.708 $mm^2$
  
  The router with Ariadne: 2.761 $mm^2$
- Slightly more overhead than Vicis
- Three times less than Immunet

<table>
<thead>
<tr>
<th></th>
<th>Wire</th>
<th>Routing Tables</th>
<th>Overhead $mm^2$</th>
<th>%Overhead</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ariadne</td>
<td>1</td>
<td>Adaptive</td>
<td>0.053</td>
<td>1.97%</td>
</tr>
<tr>
<td>Vicis</td>
<td>1</td>
<td>Deterministic</td>
<td>0.040</td>
<td>1.48%</td>
</tr>
<tr>
<td>Immunet</td>
<td>4</td>
<td>Adaptive, deterministic and a small safe table</td>
<td>0.162</td>
<td>5.98%</td>
</tr>
</tbody>
</table>
Summary

- Deadlock Avoidance
  - Normal Operation: up*/down*
  - Reconfiguration
    - Eject and re-inject
    - Cost: an additional dedicated flitbuffer per NIC

- On-Chip
  - A fully “distributed” solution
  - Require a global clock to support atomicity

- Protect ARIADNE
  - Triple Modular Redundancy (TMR)

- Partitioned Network
  - The packet delivery is guaranteed within each partition
  - Cannot transfer packets among partitions
Conclusion

- High performance and deadlock-free routing by utilizing up*/down*
  - Performance gain: 40% to 140% at 50 faults during normal operation

- High reliability
  - If a path between two nodes exists, at least one deadlock-free path is guaranteed

- A fully distributed solution with simple hardware and low complexity
  - Nodes coordinates to explore the surviving topology
  - 1.97% area overhead
Discussion

- Are the two benchmarks representative enough?
- Does the packet size affect the result?
- Does the topology influence the result?
- How about the scalability?
- What is the power consumption during reconfiguration?