Modeling the impact of interconnect technology on CMP architecture performance
RF-Interconnect as a test case

Eran Socher
School of Electrical Engineering
Tel Aviv University

4th Multi Core Forum
June 23rd 2009
Outline

• The scaling challenge in chip-multi-processors
• The RF-Interconnect approach and implementation
• MORFIC architecture and modeling
• MORFIC performance impact: run-cycles and power, NoC adaptivity
• Conclusions
Current Trend in CMP - NoC

- 65nm CMOS 80 tile NoC
- 10X8 2D mesh network-on-chip running @ 4GHz
- Bisection bandwidth 256GB/s
- 1 TFLOPS @ 1V about 98W

* An 80-Tile 1.28TFLOPS Network-on-Chip in 65nm CMOS, Vangal et al., Intel, ISSCC 2007
What is The Challenge?

- Cores would keep shrinking in size but maintain the same operation frequency (2~4GHz) due to thermal constraints.
- More cores would be integrated on the same chip to achieve performance boost through parallelism.
- Performance would be limited by the communication efficiency between cores and memories on- and off-chip.
The Scaling Trend

- Scaling reduces delay of logic gates but not wires

Transistor and Wire Delay Trend in CMOS

- Delay [ps]
- Technology Node
- 180nm, 130nm, 90nm, 65nm, 45nm, 32nm

- FO4
- 1mm RC global wire
- Repeated 1mm RC global wire
Traditional Interconnect

- Units communicate through a parallel bus using voltage signaling (charging and discharging the wire capacitance)
- Latency is RC limited (~$L^2$)
- Using CMOS repeaters reduces latency (~$L$) but does not benefit from scaling
- Supply no longer scales due to leakage
- Baseband-only signaling requires extensive equalization
- Waste of most of the bandwidth available from modern CMOS devices ($f_t > 150$GHz, $f_{\text{max}} > 250$GHz)
Major Interconnect Issues

- Latency is large across chip
- Bandwidth is RC limited (~1Gbps/wire)
- Communication pattern is fixed
- Energy consumption is high and not scalable (~10pJ/bit)
- Future microprocessors may encounter communication congestion and most of the energy will be spent on “talking” instead of computing
How Can RF Help?

- EM waves travel at the (effective) speed of light (~10ps/mm)
- Carrier frequencies can be modulated by modern CMOS with high data rates
- Transmission lines on- or off-chip can guide the waves (RF modulated data) from the transmitter to receiver with recoverable attenuation
• Data transmit through transmission lines at the speed of light, with less dispersion across the band and less baseband interference
• data rate is only limited by CMOS mixer modulation speed

* Socher & Chang, 2007 IEEE Communication Magazine
RF-I using Multi-band FDMA

• More bands are used with same modulation speed at each band
• Higher aggregate data rates can be achieved on the same transmission line
Tri-band ASK Modulated RF-I

- Transformers are used for coupling
- No carrier generation is needed in the receiver
- RF bands use differential mode and baseband uses the common mode in the TL
Schematics of the ASK Tx

- Transformer is used for isolation between VCO core and the modulator to reduce cross-band modulation
- Simple on-off modulation is used
- Second transformer is used for resonant magnetic coupling into the TL
Differential Transmission Line

- Loss of 0.6-1.6 dB/mm
Schematics of the ASK Rx

- A transformer is used for resonant coupling between the TL and each RF band receiver.
- A self-mixer is used to down-convert the data into baseband.
- The transmitted baseband data in the common mode is tapped in the connection point of the two RF transformers.
Equivalent Baseband Tx and Rx

- Capacitive coupling is used to couple the baseband data into the TL in the common mode using a low voltage swing
- The common mode is tapped at the receiver and recovered using a buffer
Die Photo

- The chip was fabricated in IBM 90nm digital process
- The TL was folded to save silicon area
- Metal fill was minimized in the TL and transformer area
Output Data Streams

- Each RF channel carries 4Gbps
- The baseband carries 2Gbps to achieve an aggregate 10Gbps data rate
- BER was measured as $<10^{-9}$ at simultaneous transmission
The output of the TL was measured using a 67GHz G-S probe.

Only the RF odd modes are measured.

Spectrum is shown for the carriers and when modulating PBRS data.
## Performance Comparison

<table>
<thead>
<tr>
<th>Interconnect Technique</th>
<th>This Work</th>
<th>Ho et. al., JSSC 2008</th>
<th>Jose et. al., JSSC 2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>Channels</td>
<td>RF-I</td>
<td>Low-Swing Capacitive Coupling</td>
<td>Low-Swing CML</td>
</tr>
<tr>
<td>Data Rate in RF Channel (Gbps)</td>
<td>4</td>
<td>NA</td>
<td>NA/NA</td>
</tr>
<tr>
<td>Data Rate in BB Channel (Gbps)</td>
<td>2</td>
<td>1</td>
<td>8, 3</td>
</tr>
<tr>
<td>Total Aggregate Data Rate (Gbps)</td>
<td>10</td>
<td>1</td>
<td>8, 3</td>
</tr>
<tr>
<td>Latency (ps/mm)</td>
<td>6</td>
<td>55</td>
<td>30, 6.5</td>
</tr>
<tr>
<td>Energy Per Bit (RF) pJ/bit/mm</td>
<td>0.09*</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Energy Per Bit (BB) pJ/bit/mm</td>
<td>0.125</td>
<td>0.28</td>
<td>0.18, 0.14</td>
</tr>
</tbody>
</table>

* Assuming shared VCO power between multiple RF-I links
Future Trends in Multi-band ASK RF-I

<table>
<thead>
<tr>
<th>Technology</th>
<th># of Carriers</th>
<th>data rate per carrier (Gb/s)</th>
<th>Total Data rate per link (Gb/s)</th>
<th>Power (mW)</th>
<th>Energy per bit(pJ)</th>
<th>Area (TX+RX) mm²</th>
<th>Area/Gbit (µm²/Gbit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>90nm</td>
<td>3RF + 1 BB</td>
<td>5</td>
<td>20</td>
<td>20</td>
<td>1.00</td>
<td>0.022</td>
<td>1100</td>
</tr>
<tr>
<td>65nm</td>
<td>4RF + 1 BB</td>
<td>6</td>
<td>30</td>
<td>25</td>
<td>0.83</td>
<td>0.0238</td>
<td>800</td>
</tr>
<tr>
<td>45nm</td>
<td>5RF + 1 BB</td>
<td>7</td>
<td>42</td>
<td>30</td>
<td>0.71</td>
<td>0.0228</td>
<td>540</td>
</tr>
<tr>
<td>32nm</td>
<td>6RF + 1 BB</td>
<td>8</td>
<td>56</td>
<td>35</td>
<td>0.63</td>
<td>0.0211</td>
<td>380</td>
</tr>
<tr>
<td>22nm</td>
<td>7RF + 1 BB</td>
<td>9</td>
<td>72</td>
<td>40</td>
<td>0.56</td>
<td>0.0193</td>
<td>260</td>
</tr>
</tbody>
</table>

Scaling in Energy per bit (pJ/bit)

Scaling in (TX+RX Area)/Gbit
Interconnect Topology Comparison

- Comparison across process technology of...
  - Traditional RC parallel bus
  - RF-Interconnect
  - Optical Interconnect
- As process technology scales toward 22nm...
  - RF-I has lowest latency
  - RF-I consumes least energy
  - RF-I has highest data rate density
- RF-I is fully compatible with modern CMOS technology
RF-Interconnect for NoC

- RF-I is built on top of 2D-Mesh NoC and serves as a “super-highway”
- Multiple carrier frequencies in the RF and MMW range (100GHz to over 500GHz)
- Data encoding by amplitude modulation of carrier
- Direct coupling between the transmission line and electronic circuits
- Improves with device performance scaling (higher data rates, more carriers)
- Potentially lower energy consumption
RF-I Physical / Logical Organization

• Physically
  – RF-I is shared bundle of transmission lines
  – Connected to and shared between set of RF-enabled routers

• Logically
  – RF-I behaves as set of N express channels
  – Each channel assigned to source, destination router pair \((s,d)\)
    • Both \(s\) and \(d\) must be RF-enabled
Architectural Challenges of RF-I

- How many/which routers should be RF-enabled?
  - How many RF-I ports should each router have?
    - Dedicated or multiplexed with other ports?
- How much RF-I bandwidth to allocate?
  - Total? Per communicating pair?
  - Impacts active layer area consumed by RF-I components
- Which routing strategy to employ in presence of RF-I express channels?
- Dynamic or static allocation of frequency bands to sources/destinations
  - Dynamic: requires arbitration overhead for channel assignment
  - Static: may miss opportunity to match changing communication demand
Our decisions...

• How many/which routers should be RF-enabled?
  – 16 routers (3 per quadrant and 4 in center)
  – How many RF-I ports should each router have?
    • 16 dedicated ports
• How much RF-I bandwidth to allocate?
  – Start with 256B total, 16B per cycle per communicating pair
• Which routing strategy to employ in presence of RF-I express channels?
  – Shortest-Path Routing
• Dynamic or static allocation of frequency bands to sources/destinations
  – Static, to save overhead
Baseline Mesh Interconnect Topology

- 10x10 mesh of 5-cycle pipelined routers
  - NoC runs at 2GHz
  - XY/YX routing
- 64 4GHz 3-wide processor cores containing
  - 8KB L1 Data Cache
  - 8KB L1 Instruction Cache
- 32 L2 Cache Banks
  - 256KB each
  - Organized as shared NUCA cache
- 4 Main Memory Interfaces
  - Labeled with + in the figure
MORFIC: Mesh Overlaid with RF-InterConnect

- Shared Z-shaped RF waveguide
- Organized as 8 bidirectional shortcut links
- Each direction of each shortcut can transmit simultaneously over shared medium
- Router A can send a flit to other router A, B to B, … H to H in a single cycle
- Router labeled X cannot directly send to any router not labeled X
  - E.g. Router B in upper left cannot send to router E in upper right directly
  - However, B in upper left can send to B in upper right, and then north to E using normal mesh link
MORFIC Results For 256B Total RF-I

- Design simulated using a modified SESC framework
- 256B RF-I consumes 0.18% silicon overhead on 400mm² die
  - RF-I components: 0.13%, Router overhead: 0.05%
- Normalized Splash-2 Execution Time and Average Packet Latency Results
  - Normalized to baseline mesh run-cycles/latency at 1
  - Average 13% (max 18%) performance improvement
  - Average 22% (max 24%) packet latency improvement
Varying Total RF-I Bandwidth

- Application performance can degrade by more than 400% for small RF-I allocations
- Too many packets waiting at RF-I access points
- RF-I shortcuts become bottlenecks
Router Activity and Congestion

- Lighter shade represents more activity
- X% usage: (100-X)% of packets are XY/YX routed (no RF-I), X% packets routed on shortest path
- Utilizing 32B RF-I 25% of the time spreads router activity while avoiding bottlenecks at shortcut-access points (compared to 100% usage)
Varying RF-I Utilization (1 / 2)

- **Search-and-Set technique**
  - Finds best utilization and locks it for rest of app execution

Drop utilization 10% and lock!
• Fully-adaptive can improve performance by as much as:
  - 6.5% over baseline on 32B RF-I allocation
  - 14.3% over baseline on 96B RF-I allocation

**NETWORK CONDITIONS CHANGE!**

If under-utilized

If over-utilized

**Diagram Description:**
- X-axis: Time (ranging from 0 to 250k steps)
- Y-axis: Utilization (%)
- Red line: Graphical representation of varying RF-I utilization over time
- The diagram illustrates the times when network conditions change and the corresponding utilization levels.
Using RF-I to Adapt an NoC

• Conventional NoCs provide same bandwidth on all paths of an application
  – In order to support all kind of traffic demands
• Observation: Communication demands tend to vary
  – Between different applications
    • “FOCUS” of this work
  – Within a single application, during different periods of execution
• We Adapt the NoC by tuning to the varying communication patterns
Benignits of RF-I

Radio Frequency – Interconnect (RF-I):
• Cutting edge transmission line technology
• Enables adaptive NoC by
  – Flexible bandwidth allocation
  – Matching the communication demands of applications
• Offers dramatic power and area savings
  – Simplification of NoC topology
• Natural means of multicast
• Basic technique:
  – Provide bandwidth only where needed
  – Reduce overall bandwidth to offer power savings
Application Diversity

• For a 100 (10x10 mesh) router configuration:
  • Measures messages sent from a router on x-axis to router on y-axis
• Legend for the figure on the coming slide
  — **Black:** no traffic
  — **Dark Blue:** $[1, \text{mean} / 4)$
  — **Light Blue:** $[\text{mean}/4, \text{mean}/2)$
  — **White:** $[\text{mean}/2, 2\times\text{mean})$
  — **Orange:** $[2\times\text{mean}, 4\times\text{mean}]$
  — **Red:** $(4\times\text{mean}, \text{inf})$
Messages Sent between Routers

LU

High communication

High communication
Baseline Architecture

- 10x10 mesh of pipelined routers
  - NoC runs at 2GHz
  - XY routing
- 64 4GHz 3-wide processor cores containing
  - 8KB L1 Data Cache
  - 8KB L1 Instruction Cache
- 32 L2 Cache Banks
  - 256KB each
  - Organized as shared NUCA cache
- 4 Main Memory Interfaces
  - Labeled with + in the figure
RF-I Physical Organization

- Physically
  - RF-I is a bundle of transmission lines
  - Connected to and shared between set of RF-enabled routers
  - RF-enabled router consists of a Tx/Rx pair
RF-Enabled Routers

RF-Enable 50 Routers - Represented by GREEN

Routing Tables

Add 6th Port

Transmission Line…
RF-I Logical Organization

• Logically:
  - RF-I behaves as set of N express channels
  - Each channel assigned to src, dest router pair \((s,d)\)

• Reconfigured by:
  - remapping shortcuts to match needs of different applications
How can RF-I Lead to Power Savings?

- We can thin the baseline mesh links
  - From 16B...
  - …to 8B
  - …to 4B

- Can RF-I make up the difference in performance while saving overall power?
  - RF-I provides bandwidth where most necessary
  - Baseline RC wires supply the rest
Architecture-Specific Shortcuts

- Design time shortcuts
- Referred to as static shortcut in the remainder of this talk
- Selection Criteria
  - Consider an optimization function for a topology
    \[ W_{x,y} = \begin{cases} 
    \text{length of shortest-path}(x,y) & \text{if } x \neq y \\
    0 & \text{if } x == y 
    \end{cases} \]
  - We wish to minimize the total cost of the graph G representing the network-on-chip
    \[ \text{Total-Cost}(G) = \sum_{\text{all}(x,y)} W_{x,y} \]
Shortcut-Selection Constraints

• Each router should have at most 6 ports
  – A router can be at most one shortcut source and at most one shortcut destination

• Total of $B$ (budget) unidirectional shortcuts: $B = 16$

• For static shortcuts:
  – RF-enable routers which are shortcut srcs/dests
  – At most 16 RF-enabled routers

• For adaptive shortcuts, shortcut srcs/dests are limited to
  – RF-enabled routers chosen at design-time
Min Total-Cost(NoC): Heuristic 1

O(BV^5)

I) For each pair of non-adjacent routers i,j
   - Make a new permutation graph G_{i,j} with an edge between them
   - Calculate Total-Cost(G_{i,j})
   - Record how much improvement as... |Total-Cost(G_{i,j}) – Total-Cost(G)|

II) Select shortcut of edge (x,y) such that G_{x,y} had max improvement
   - Disallow any use of x as a src or y as a dest hence

III) Repeat (I) and (II) until budget B exhausted
Min Total-Cost(NoC): Heuristic 2

O(BV^3)

(I) Calculate \( W_{i,j} \) for all pairs \( i,j \) in \( G \)
   - Record all \( W_{i,j} \)

(II) Select shortcut of edge \((x,y)\) s.t
   \[ W_{x,y} = \max(W_{i,j}) \]
   - Disallow any use of \( x \) as a src or \( y \) as a dest hence

(III) Repeat (I) and (II) until budget exhausted

These shortcuts tend to perform within 1% as well as those chosen with heuristic 1
Adaptive RF-I Shortcuts

• Assume a profile of communication for an application
  • \( F_{i,j} \) = count of messages sent between router \( i \) and router \( j \)

• Change optimization function
  \[
  \text{Total-Cost}(G) = \sum_{\text{all}(x,y)} (F_{x,y} \cdot W_{x,y})
  \]

• To offset effect of removing src/dest routers (already selected) from consideration
  • Alternate router-to-router shortcuts with region-to-region shortcuts
  • Allows placement of shortcuts at routers near a hotspot
Evaluation Methodology

• Use detailed interconnection network simulator - Garnet[1]
• Build probabilistic traces:
  – To cover different communication patterns
  – Maybe exhibited by future applications
• Leverage Orion[4], CosiNoC[5], IPEM[2] for power methodology
RF-I Enables Power Savings

- On average **adaptive** shortcuts w/ 50 RF-enabled routers on a 4B mesh
  - 62%(82%) power(area) savings over baseline
  - Performance comparable to baseline
Unified Analysis

- Adaptive RF-I enabled NoC
  - Cost Effective in terms of both power and performance
Summary

• RF-I is a potential candidate to alleviate the communication bottleneck of future chip multiprocessors.

• Multi-band ASK-based communication on-chip was demonstrated at mm-wave frequencies and baseband achieving 10Gbps and down to 0.09pJ/bit/mm and 6ps/mm latency at BER<10^{-9}.

• The scaling trend benefits multi-band RF-I as opposed to alternative approaches.
Summary

• We introduce RF-I technology for on-chip communication

• We present MORFIC architecture
  – 64 Cores, 32 L2 Cache Banks, 4 Memory Interfaces
  – RF-I provides an average 13% (max 18%) performance improvement for area cost of 0.18% of active layer

• RF-I access points may become bottlenecks
  – Adapting RF-I utilization to changing network conditions can avoid congestion at access points
Summary

RF-I:

• Enables adaptive NoC
  – Bandwidth can be flexibly allocated
  – To match the communication demands of applications

• Offers dramatic power and area savings
  – By simplifying baseline NoC topology
  – Provides performance of 16B mesh on a 4B mesh
    • 62% power savings, 82% area savings

• Natural means of multicast
Thank You!

• Acknowledgments:
  M-C. Frank Chang, Jason Cong, Adam Kaplan, Chunyue Liu, Mishali Naik, Jagannath Premkumar, Glenn Reinman, Sai-Wang Tam – EE & CS UCLA