On-Chip Interconnects: Circuits and Signaling from an MPSoC Perspective

Wayne Burleson
Associate Professor, Dept ECE
University of Massachusetts Amherst

burleson@ecs.umass.edu
My Perspective

- VLSI Signal Processing, BS/MS MIT 1983, PhD Colorado 1989
- Worked as a VLSI Designer (Fairchild, VTI) and teach VLSI Design
- Research in VLSI Circuits
  - Low-power (NSF, SRC)
  - Interconnects (SRC, Intel)
  - Wave-pipelining (NSF)
  - SRAM (Intel, CRL)
  - Soft-errors (Intel, MMDC)
- Research in VLSI Architecture
  - Adaptive SOC (NSF)
  - VLSI Signal Processing (NSF)
  - Video, 3D Graphics (NSF)
  - Embedded Security (HSARPA)
Objectives

- Provide insight into on-chip interconnects from a circuit designer’s perspective
- Survey recent research in on-chip interconnects
- Present MPSoC interconnect requirements and metrics
- Show how to compare circuit and signaling solutions
- Discuss the impact of uncertainties (process variation, noise, temperature)
- Show CAD support for interconnect design and estimation
- Some examples
On-Chip Interconnect: Levels of Abstraction

- **Network level**
  - CDMA
  - TDMA

- **System level**
  - Communication Links
  - Adaptive supply voltage links

- **Architecture level**
  - AMBA™
  - CoreConnect™

- **Circuit level**
  - Low Swing
  - Coding
  - Single / Differential
MPSoC Interconnect Requirements

- Intra-core vs. inter-core,
- Bus-width (1, 8, …, 32, 64, …)
- Adjacent core vs. long-haul
- Repeated vs. unrepeated
- Single-cycle vs. pipelined
- Bus vs. point-to-point
- Synch vs. asynch

Metrics:
- Latency
- Bandwidth
- Noise
- Area
- Power/Energy

MPSoC NoC Interconnect issues

- Granularity of cores, i.e., wirelength
- Bus width
- Network topology
- Link layer (fault-tolerance, etc.)
- Synchronous vs. Asynchronous

MPSoC “Bus” Alternatives

- **Fixed Bus** [Bergamaschi, DAC, 2000]
  - Point to point communication
  - Signals between cores transferred by dedicated wires

- **FPGA-like Bus** [Cherepacha, FPGA Sym, 1994]
  - Programmable interconnects
  - Employ static network

- **Arbitrated Bus** [IDT Inc., 2000]
  - Time-shared multiple core connectivity
  - Use arbitrator

- **Hierarchical Bus** [AMBA, ARM Inc]
  - Combine multiple buses using bus bridges
  - Separate buses for cores and I/O

- **NoC Bus** [Dally, DAC, 2000]
  - Resources communicate with data packets
  - Use switch fabric

SoC Bus Standards: CoreConnect™ and AMBA™

<table>
<thead>
<tr>
<th></th>
<th>IBM CoreConnect Processor Local Bus</th>
<th>ARM AMBA 2.0 AMBA High-performance Bus</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bus Architecture</strong></td>
<td>32-, 64-, and 128-bits, extendable to 256-bits</td>
<td>32-, 64-, and 128-bits</td>
</tr>
<tr>
<td><strong>Data Buses</strong></td>
<td>Separate Read and Write</td>
<td>Separate Read and Write</td>
</tr>
<tr>
<td><strong>Key Capabilities</strong></td>
<td>Multiple Bus Masters</td>
<td>Multiple Bus Masters</td>
</tr>
<tr>
<td></td>
<td>4 Deep Read Pipelining</td>
<td>Pipelining</td>
</tr>
<tr>
<td></td>
<td>2 Deep Write Pipelining</td>
<td>Split Transactions</td>
</tr>
<tr>
<td></td>
<td>Split Transactions</td>
<td>Burst Transfers</td>
</tr>
<tr>
<td></td>
<td>Burst Transfers</td>
<td>Line Transfers</td>
</tr>
<tr>
<td>On-Chip Peripheral Bus</td>
<td>Supports Multiple Masters</td>
<td>Single Master: The APB Bridge</td>
</tr>
<tr>
<td></td>
<td>Data Buses: Separate Read and Write</td>
<td>Separate or 3-state</td>
</tr>
</tbody>
</table>

IBM, 2000
Interconnect Geometry Scaling

- Weak scaling of vertical dimension compared to horizontal dimension
- Extremely high height/width aspect ratios
- Reduces degradation of interconnect resistance

<table>
<thead>
<tr>
<th>Layer</th>
<th>Pitch (nm)</th>
<th>Thick (nm)</th>
<th>AspectRatio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isolation</td>
<td>220</td>
<td>320</td>
<td>-</td>
</tr>
<tr>
<td>Polysilicon</td>
<td>220</td>
<td>90</td>
<td>-</td>
</tr>
<tr>
<td>Contacted gate pitch</td>
<td>220</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Metal 1</td>
<td>210</td>
<td>170</td>
<td>1.6</td>
</tr>
<tr>
<td>Metal 2</td>
<td>210</td>
<td>190</td>
<td>1.8</td>
</tr>
<tr>
<td>Metal 3</td>
<td>220</td>
<td>200</td>
<td>1.8</td>
</tr>
<tr>
<td>Metal 4</td>
<td>280</td>
<td>250</td>
<td>1.8</td>
</tr>
<tr>
<td>Metal 5</td>
<td>330</td>
<td>300</td>
<td>1.8</td>
</tr>
<tr>
<td>Metal 6</td>
<td>480</td>
<td>430</td>
<td>1.8</td>
</tr>
<tr>
<td>Metal 7</td>
<td>720</td>
<td>650</td>
<td>1.8</td>
</tr>
<tr>
<td>Metal 8</td>
<td>1080</td>
<td>975</td>
<td>1.8</td>
</tr>
</tbody>
</table>

65nm Intel® Technology

90nm Intel® Technology

Interconnect RC Trend

<table>
<thead>
<tr>
<th>Layer</th>
<th>R</th>
<th>C</th>
<th>RC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poly</td>
<td>45%</td>
<td>-2%</td>
<td>42%</td>
</tr>
<tr>
<td>M1</td>
<td>53%</td>
<td>5%</td>
<td>61%</td>
</tr>
<tr>
<td>M2</td>
<td>46%</td>
<td>12%</td>
<td>62%</td>
</tr>
<tr>
<td>M3</td>
<td>39%</td>
<td>8%</td>
<td>51%</td>
</tr>
<tr>
<td>M4</td>
<td>18%</td>
<td>24%</td>
<td>46%</td>
</tr>
</tbody>
</table>

- RC/µm increases 40-60% per generation
- Copper, low-K dielectric: modest benefit
Interconnect Distribution Trend

RC/\mu m scaling trend is only one side of the story…
Average wire lengths don’t scale well
Interconnect Power Consumption

- Using Vdd programmability
- High Vdd to devices on critical path
- Low Vdd to devices on non-critical paths
- Vdd Off for inactive paths

A – Baseline Fabric
B – Fabric with Vdd Configurable Interconnect

This work builds on a similar idea for FPGAs described in:
Interconnect Modeling

a. Capacitive model
b. RC model
c. RLC model

Massoud, Ckts and Devices Mag, 2001
Interconnect Issues – Signal Integrity

- Inductance becoming important
- Self inductance results in ringing
- Mutual inductance results in crosstalk

Ismail, TVLSI, 2002
Circuit and Signaling Solutions

- Conventional Circuit techniques
  - Repeater insertion
  - Booster insertion

- Low Swing techniques
  - Pseudo differential interconnect
  - Differential Current sensing

- Bus encoding techniques
  - Transition aware encoding
  - Low Power encoding for crosstalk reduction

- Signaling techniques
  - Multi-level signaling
  - Near speed of light signaling
Interconnect Circuits - Repeater

- Optimum repeater insertion reduces interconnect delay.
- Optimized energy-delay tradeoffs used to satisfy design criteria.
Interconnect Circuits - Repeaters

Dynamic Power in Repeated Wires

Magen, SLIP, 2004

Repeaters per mm

Technology (nm)

Repeaters per CPU Clock Cycle

180 130 100 70 45

0 0.5 1 1.5 2 2.5 3 3.5

Maheshwari, PhD Thesis, 2004

Gate 34%

Interconnect 51%

Diffusion 15%

Dynamic Power in Repeated Wires

Magen, SLIP, 2004

# of Repeaters

Itanium, Power4 range

10^4 10^5 10^6 10^7

Rent's Exponent

180 nm 130 nm 100 nm 70 nm

Kaul, TVLSI, 2004
Interconnect Circuits – Boosters

Nalamalpu, TrCAD, 2002, Kaul, TVLSI, 2004
Interconnect Circuits - Low swing

- Reduce the swing on the interconnect
- Use a PMOS/NMOS device to provide a resistive path
- Reduces dynamic power since interconnect is not charged full rail.
- Noise immunity low at the output due to reduced swing
- Static power dissipation due to low impedance path

- Current mode, voltage mode
- Single ended, differential
Interconnect Circuits – Low Swing

- Voltage mode
- Single ended
- One wire per bit
- Receiver not sensitive to supply variation

Pseudo Differential Interconnect

<table>
<thead>
<tr>
<th>Schemes</th>
<th>Energy (PJ)</th>
<th>Delay (ns)</th>
<th>E+D (PJ·ns)</th>
<th>Swing (V)</th>
<th>Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Driver/Wire</td>
<td>Receiver</td>
<td>Total</td>
<td>Driver/Wire</td>
<td>Receiver</td>
</tr>
<tr>
<td>CMOS</td>
<td>11.45</td>
<td>0.15</td>
<td>11.6</td>
<td>1.64</td>
<td>0.47</td>
</tr>
<tr>
<td>PDIFF</td>
<td>1.32</td>
<td>0.60</td>
<td>1.92</td>
<td>1.65</td>
<td>0.75</td>
</tr>
</tbody>
</table>

Zhang, TVLSI, 2000
Interconnect Circuits – Low Swing

- Current mode, Differential
- Avoids charging and discharging wire capacitance
- No repeaters along the wire: Avoids placement constraints
- Suffers from static power dissipation (paths shown by dashed lines)
Delay-Power tradeoffs

Differential Current-Sensing

Maheshwari, TVLSI, 2004
Delay vs. wirelength

![Graph showing delay vs. wirelength for Intel 90nm (wires with 2x min. width). The graph compares repeaters and current-sensing, indicating a 20 – 35% gain.]
% chip coverage in n cycles

Percentage of Chip Coverage

Maheshwari, ASIC SOC, 2002
Uniform Repeater Insertion

Current-sensing e.g. DCS

\[ \sum_{i=1}^{n} L_{Ri} + L_C = L \]

How much wire driven by repeaters?

\[ \sum_{i=1}^{n} \frac{L_{Ri}}{L} = ? \]
Eliminating bus static power dissipation

- Send current only when there is a transition
- Hold the bus at GND otherwise
- Encoder and decoder overhead
Interconnect Solutions - Bus Encoding

• Reduce dynamic power due to switching activity on a bus
  – Transition encoding, spatial encoding, invert encoding, pattern encoding

• Various encoding target different aspect of interconnect
  – Delay, power, energy, crosstalk, area

• Cost of encoding/decoding
  – Power, area, latency, additional wires
Interconnect Solutions – Bus encoding

- Uses a dynamic bus configuration
- Encoder translates input transition activity into an output logic state
- Decoder uses encoded signal to reconstruct the original input using its stored state information to distinguish between the two input transitions.

Transition Encoded Dynamic Bus

Anders, JSSC, 2003
Interconnect Solutions - Bus Encoding

- Bus invert encoding
  - Checks each cycle if there is a possibility of greater than 50% transitions on the bus
  - Decides whether sending the true or compliment form of the signals
  - Reduces the switching activity
  - Requires one additional wire to inform receiver whether the bus is true or complement
  - Numerous extensions and improvements for different statistical assumptions and metrics
Multi-level Current Signaling

- Encode two or more data bits and transmit on interconnect.
- The two or more data bits are encoded into four or more current levels. Current provides more head-room than voltage!
- Sense the current levels and decode the original signals

Venkatraman, ISQED, 2005
Phase Coding

- Actually phase modulation
- Transmitting multiple bits in one transition
  - Significant power and area savings
  - Increased bandwidth
- Phase coding – Phase determines the data
- How to deal with timing uncertainty?
Open Loop Phase Coding

- Delay elements can be shared across wires
- Supply noise, Process variation etc. can result in errors
Measured Results: Closed Loop

- 16-bit 5mm long bus, 0.27u wide, 0.27u spacing, shielded, 1GHz
- Repeater insertion, Transition encoding used
- Encode in ½ cycle and use ½ cycle for decode

<table>
<thead>
<tr>
<th>Encoding Levels (bits/wire)</th>
<th>Encoder Overhead (mW)</th>
<th>Decoder Overhead (mW)</th>
<th>Phase coding power (mW)</th>
<th>Repeater bus (mW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.33</td>
<td>1.00</td>
<td>5.61</td>
<td>8.56</td>
</tr>
<tr>
<td>3</td>
<td>0.47</td>
<td>1.33</td>
<td>5.01</td>
<td>8.56</td>
</tr>
<tr>
<td>4</td>
<td>0.62</td>
<td>1.52</td>
<td>4.28</td>
<td>8.56</td>
</tr>
</tbody>
</table>
Near Speed of light Signaling

- 283ps for 20mm 16um wide AL wire in 0.18um CMOS tech
- Very wide, R~0
- Uses frequency modulation

<table>
<thead>
<tr>
<th>Signaling</th>
<th>Propagation Medium</th>
<th>Time of Flight for 20mm [ps]</th>
<th>Delay for 20 mm [ps]</th>
<th>Power [mW]</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modulation (This work)</td>
<td>Wide metal wires ((c_s = 4))</td>
<td>133</td>
<td>300</td>
<td>16</td>
<td>Large metal area</td>
</tr>
<tr>
<td>Repeateres</td>
<td>Min. sized metal wires ((c_s = 4))</td>
<td>133</td>
<td>1400</td>
<td>30</td>
<td>Slow</td>
</tr>
<tr>
<td></td>
<td>Wide metal wires ((c_s = 4))</td>
<td>133</td>
<td>400</td>
<td>50</td>
<td>Large metal area, High power</td>
</tr>
<tr>
<td>Optics (Edge-Emitting) [11]</td>
<td>Air ((c_s = 4))</td>
<td>66</td>
<td>300</td>
<td>80</td>
<td>Packaging, Integration issues</td>
</tr>
<tr>
<td></td>
<td>On-Chip Waveguide ((e_s = 11.7))</td>
<td>228</td>
<td>500</td>
<td>80</td>
<td>Integration issues</td>
</tr>
<tr>
<td>Optics (VCSEL) [11]</td>
<td>Air ((c_s = 1))</td>
<td>66</td>
<td>400</td>
<td>60</td>
<td>Packaging, Integration issues</td>
</tr>
<tr>
<td></td>
<td>On-Chip Waveguide ((e_s = 11.7))</td>
<td>228</td>
<td>600</td>
<td>60</td>
<td>Integration issues</td>
</tr>
</tbody>
</table>

Chang, JSSC, 2003
Uncertainty - Process Variations

Nassif, ISQED, 2000
Impact of Uncertainty on Delay and Power

- 100nm technology
- 1000 Monte Carlo Runs
- Power variability of 43.64%
- Delay variability of 28.95%

- 100nm Technology
- Bin 1 (High Performance) Yield – 36.1%
- Bin 2 (Low Delay) Yield – 27.3%
- Bin 3 (Low Power) Yield – 25.1%
- Bin 4 (Low Performance) Yield – 11.5%
Uncertainty - Temperature Variations

- Leakage power significantly increases with temp. in 45 nm node.
  - parabolic curvature ($y = Cx^2$) in terms of varying temperatures.
- Higher temperature sensitivity (45 nm) on delay, power and leakage.
  - Accurate RLC modelization to provide underestimation.
Uncertainty due to Power Supply Noise

\[ \sum_{k=1}^{K} \tau_k(t), \text{ where } K = 3 \]
Uncertainty – Variation-tolerant Design

- Razor methodology
  - A voltage-scaling methodology based on real-time detection and correction of circuit timing errors
  - Allows for energy tuning of microprocessor pipeline
  - Application or Razor methodology results in up to 64% energy savings with less than 3% delay penalty for error recovery

Austin, Computer Magazine, 2004
The Delay, Energy, Error space

- Low-energy, High-error, Fast (e.g. low Vdd, low Vth, no checkers)
- Slow, but highly reliable and low-energy
- Fast, reliable, more energy-hungry
- Delay Constraint (iso-delay Curve)
- Energy
- Error

Burleson, BARC 2005
Interconnect test chips

- Repeaters
- Phase coding (Open Loop)
- Phase coding (Closed Loop)
- Current-pulse Signaling
- Differential Current-sensing
- Hybrid Circuit
- Noise

Dimensions:
- 2.7 mm
- 3.3 mm
CAD Support

• GTX (SRC-MARCO)
  – A GSRC Technology Extrapolation System

• NoCIC (UMASS)
  – SPICE-based Interconnect Calculator for
    aggressive Circuit techniques (current-
    sensing, multi-bit sensing, boosters, etc.)
Repeater Optimization using GTX

- Most commonly cited optimal buffer sizing expression (Bakoglu)
- In GTX:
  - Sweep repeater size for single stage in the chain
  - Examine both delay and energy-delay product

\[ L_{seg} = 2.14 \text{ mm} \]
\[ W = S = 1 \text{ mm} \]
\[ W = S = 0.5 \text{ mm} \]

Critical Path Delay (ns)

Normalized Energy-Delay Product

Stroobandt GSRC 2000, Cao, TVLSI, 2003
Inductance analysis using GTX

- Five different models implemented in GTX
  - Bakoglu’s model (RC_B)
  - [Alpert, Devgan and Kashyap, ISPD 2000] (RC_ADK)
  - [Ismail, Friedman and Neves, TCAD 19(1), 2000] (RLC_IFN)
  - [Kahng and Muddu, TCAD 1997] (RLC_KM)
  - Extension of [Alpert, Devgan and Kashyap, ISPD 2000] (RLC_ADK)

Stroobandt GSRC 2000, Cao, TVLSI, 2003
A snapshot of NoCIC

NoCIC : Network-on-Chip Interconnect Calculator

PARAMETER SELECTION WINDOW

NoC Parameters
- Title Size: 4mm
- Bus Size: 8 Bit
- Supply Voltage: 1.8V

Interconnect Parameters:
- Signaling Technique: Manchester
- Technology: 180nm
- Clocking: Unclocked

Analysis
- Output: Delay
- Compare: Repeater
- Booster
- Differential Current Sensing
- Multi-Level Current Signaling

Submit | Reset

PLOT DISPLAY WINDOW

Venkatraman, SLIP, 2003
Case Study: MPSoC with NOC

Liu’s* Prediction
Max Core Freq: \textbf{2Ghz}(100nm)
Cost Perf: 3.5mm core
High Perf: 7.1mm core

Traditional Interconnect estimate
Assuming Repeated Wires
Int Freq: \textbf{1.7Ghz}
(interconnect limited)

NoCIC
Assuming Differential current sensing
Int Freq: \textbf{2.08Ghz}
Energy penalty: < 5%

* J. Liu et al. System level interconnect design for network-on-chip interconnect IPs, in proceedings of the international workshop on System level interconnect prediction, SLIP 2003.
Case-Study: On-Chip Security
(Burleson, Tessier, Gong, Wolf, Gogniat, 2005)

- On-chip monitoring and security bus
- Latency-critical for fast detection and mitigation of attacks
- Improved power, performance and security over software-based defenses

CM = Configurable Monitor
OCIN = On-Chip Intelligence Network
Conclusions & Challenges

- Interconnects are a critical enabling abstraction in MPSoC
- Interconnects play a very large and increasing role in delay, energy, and design effort.
- Interconnect can be solved simultaneously at the micro-architectural, circuit and process levels
- Aggressive circuit and signaling techniques show promise with minimal architectural impact
- CAD support needed, especially
  - early estimation for architecture and floorplanning
  - final verification in the presence of uncertainties
**VLSI Interconnects: A Design Perspective,**  
W. Burleson and A. Maheshwari  
Morgan-Kaufmann. 2005(6)  
• 400-page textbook with HW problems, covering:  
  – History (both off-chip and on-chip)  
  – Process (metallization, dielectrics, etc.)  
  – Architecture (processor, ASIC, FPGA, memory)  
  – Theoretical models (graph, information-theoretic)  
  – Wire models (R,C,L,M,...)  
  – Circuits (repeaters, boosters, sense-amps, etc.)  
  – CAD (estimation, synthesis, optimization)  
  – Case Studies (buses, memories, ASIC, FPGA)  
  – Future (nano, optical, wireless, etc.)
UMASS Interconnect Circuit Design Group

• Students:
  – Vishak Venkatraman (internship at CRL, PhD 06)
  – Jinwook Jang (MS 05, PhD 08)
  – Sheng Xu (new Sept 04)
  – Ibis Benito (new Jan 05)
  – Atul Maheshwari (now at Intel)
  – Matt Heath (now at Intel)
  – Aiyappan Natarajan (now at AMD)
  – Vijay Shankar (now at Qualcomm)
  – Anki Nalamalpu (now at Intel)

• Collaborators:
  – Sandip Kundu (UMASS/Intel/IBM)
  – Russ Tessier, Israel Koren, Aura Ganz (UMASS)
  – Shubu Mukherjee, Rich Watson, (Intel MMDC)
  – SRC liaisons (Intel CRL, Freescale)

• Alums
  – Manoj Sinha (now at Micron)
  – Chris Cowell (now at Intel)
  – Sriram Srinivasan (now at AMD)
  – Andrew Laffely (now Prof at USAFA)
  – Srividya Srinivasaragavan (now at Intel)