Chapter 6
Optimizing Power @ Design Time – Interconnect and Clocks

Slide 6.1

Optimizing Power @ Design Time

Interconnect and Clocks

Jan M. Rabaey
Chapter Outline

- Trends and bounds
- An OSI approach to interconnect-optimization
  - Physical layer
  - Data link and MAC
  - Network
  - Application
- Clock distribution

Slide 6.2
### ITRS Projections

<table>
<thead>
<tr>
<th>Calendar Year</th>
<th>2012</th>
<th>2018</th>
<th>2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interconnect One Half Pitch</td>
<td>35 nm</td>
<td>18 nm</td>
<td>14 nm</td>
</tr>
<tr>
<td>MOSFET Physical Gate Length</td>
<td>14 nm</td>
<td>7 nm</td>
<td>6 nm</td>
</tr>
<tr>
<td>Number of Interconnect Levels</td>
<td>12–16</td>
<td>14–18</td>
<td>14–18</td>
</tr>
<tr>
<td>On-Chip Local Clock</td>
<td>20 GHz</td>
<td>53 GHz</td>
<td>73 GHz</td>
</tr>
<tr>
<td>Chip-to-Board Clock</td>
<td>15 GHz</td>
<td>56 GHz</td>
<td>89 GHz</td>
</tr>
<tr>
<td># of Hi Perf. ASIC Signal I/O Pads</td>
<td>2500</td>
<td>3100</td>
<td>3100</td>
</tr>
<tr>
<td># of Hi Perf. ASIC Power/Ground Pads</td>
<td>2500</td>
<td>3100</td>
<td>3100</td>
</tr>
<tr>
<td>Supply Voltage</td>
<td>0.7–0.9 V</td>
<td>0.5–0.7 V</td>
<td>0.5–0.7 V</td>
</tr>
<tr>
<td>Supply Current</td>
<td>283–220 A</td>
<td>396–283 A</td>
<td>396–283 A</td>
</tr>
</tbody>
</table>

*Slide 6.3*
Increasing Impact of Interconnect

- Interconnect is now exceeding transistors in
  - Latency
  - Power dissipation
  - Manufacturing complexity
- Direct consequence of scaling
Communication Dominant Part of Power Budget

- Control: 15%
- I/O Drivers: 10%
- Execution Units: 40%
- Caches: 20%
- Clocks: 15%
- Interconnect: 65%

Slide 6.5
### Idealized Wire Scaling Model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Relation</th>
<th>Local Wire</th>
<th>Constant Length</th>
<th>Global Wire</th>
</tr>
</thead>
<tbody>
<tr>
<td>W, H, t</td>
<td>1/S</td>
<td>1/S</td>
<td>1/S</td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>1/S</td>
<td>1</td>
<td>1/S</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>LW/Ht</td>
<td>1/S</td>
<td>1</td>
<td>1/S</td>
</tr>
<tr>
<td>R</td>
<td>L/WH</td>
<td>S</td>
<td>S^2</td>
<td>S^3/S C</td>
</tr>
<tr>
<td>C - CR</td>
<td>L/WHt</td>
<td>1</td>
<td>S^2</td>
<td>S^3/S C^2</td>
</tr>
<tr>
<td>E</td>
<td>CV^2</td>
<td>1/SU^2</td>
<td>1/U^2</td>
<td>1/(S U^2)</td>
</tr>
</tbody>
</table>
Distribution of Wire Lengths on Chip

Interconnect Density Function $f(t)$

- Actual Data
- Stochastic Model

$N = 86161$
$p = 0.8$
$k = 5.0$

Interconnect Length, $\Gamma$ [gate pitches]

[Ref: J. Davis, C&S 98]
Technology Innovations

Reduce resistivity (e.g., Copper)

Reduce dielectric permittivity (e.g., Aerogels or air)

Reduce wire lengths through 3D-integration

Novel interconnect media (carbon nanotubes, optical)

(Courtesy of IBM and IFC FCRP)
Logic Scaling

$P_t \sim 1/s^3$

[Ref: J. Davis, Proc'01]

Slide 6.9
Interconnect Scaling

\[ L^{-2} \tau = 10^{-5} \text{[s/cm}^{-2}] \]

\[ L^{-2} \tau \sim S^2 \]

[Ref: J. Davis, Proc'01]
Lower Bounds on Interconnect Energy

Shannon's theorem on maximum capacity of communication channel

\[ C \leq B \log_2 \left(1 + \frac{P_s}{kT} \right) \]

\[ E_{\text{bit}} = \frac{P_s}{C} \frac{kT}{B} \]

\( C \): capacity in bits/sec
\( B \): bandwidth
\( P_s \): average signal power

\[ E_{\text{bit}} \text{ (min)} = E_{\text{bit}} \frac{(C / B \to 0) = kT \ln(2)}{C/B} \]

Valid for an “infinitely long” bit transition \((C/B \to 0)\)
Equals \(4.10^{-21}\) J/bit at room temperature

[Ref: J. Davis, Proc’01]
Reducing Interconnect Power/Energy

- Same philosophy as with logic: reduce capacitance, voltage (or voltage swing), and/or activity
- A major difference: sending a bit(s) from one point to another is fundamentally a communications/networking problem, and it helps to consider it as such
- Abstraction layers are different:
  - For computation: device, gate, logic, micro-architecture
  - For communication: wire, link, network, transport
- Helps to organize along abstraction layers, well-understood in the networking world: the OSI protocol stack
OSI Protocol Stack

- Reference model for wired and wireless protocol design
  — Also useful guide for conception and optimization of on-chip communication
- Layered approach allows for orthogonalization of concerns and decomposition of constraints
- No requirement to implement all layers of the stack
- Layered structure need not necessarily be maintained in final implementation

[Ref: M. Sgroi, DAC’01]
The Physical Layer

- Transmits bits over physical interconnect medium (wire)
  - Physical medium
    - Material choice, repeater insertion
  - Signal waveform
    - Discrete levels, pulses, modulated sinusoids
  - Voltages
    - Reduced swing
  - Timing, synchronization

So far, on-chip communication almost uniquely “level-based”
Repeater Insertion

**Optimal receiver insertion** results in wire delay linear with $L$

$$t_p \propto L \sqrt{(R_d C_d)(r_w c_w)}$$

with $R_d C_d$ and $r_w c_w$ intrinsic delays of inverter and wire, respectively

**But:** At major energy cost!
Repeater Insertion — Example

- 1 cm Cu wire in 90 nm technology (on intermediate layers)
  - \( r_w = 250 \, \Omega/mm; \ c_w = 200 \, fF/mm \)
  - \( t_p = 0.69r_wc_wL^2 = 3.45 \, ns \)

- Optimal driver insertion:
  - \( t_{p,\text{opt}} = 0.5 \, ns \)
  - Requires insertion of 13 repeaters
  - Energy per transition 8 times larger than just charging the wire (6 pJ versus 0.75 pJ!)

- It pays to back off!
Wire Energy–Delay Trade-off

L = 1 cm (Cu) 90 nm CMOS

Repeat overhead

wire energy only

Slide 6.17
Multi-Dimensional Optimization

- Design parameters: Voltage, number of stages, buffer sizes
- Voltage scaling has the largest impact, followed by selection of number of repeaters
- Transistor sizing secondary
Reduced Swing

- $E_{\text{bit}} = CV_{DD}V_{\text{swing}}$
- Concerns:
  - Overhead (area, delay)
  - Robustness (supply noise, crosstalk, process variations)
  - Repeaters?
Traditional Level Converter

- Requires two discrete voltage levels
- Asynchronous level conversion adds extra delay

[Ref: H. Zhang, TVLSI'00]
Avoiding Extra References

[Ref: H. Zhang, VLSI’00]

Slide 6.21
Differential (Clocked) Signaling

- Allows for very low swings (200 mV)
- Robust
- Quadratic energy savings
- But: doubling the wiring, extra clock signal, complexity

[Ref: T. Burd, UCB’01]
Lower Bound on Signal Swing?

- Reduction of signal swing translates into higher power dissipation in receiver – trade-off between wire and receiver energy dissipation
- Reduced SNR impacts reliability – current on-chip interconnect strategies require Bit Error Rate (BER) of zero (in contrast to communication and network links)
  - Noise sources: power supply noise, crosstalk
- Swings as low as 200 mV have been reported [Ref: Burd’00], 100 mV definitely possible
- Further reduction requires crosstalk suppression
Quasi-Adiabatic Charging

- Uses stepwise approximation of adiabatic (dis)charging
- Capacitors acting as “charge reservoir”
- Energy drawn from supply reduced by factor $n$

[Ref: L. Svensson, ISLPED’96]
Charge Redistribution Schemes

- Charge recycled from top to bottom
- Precharge phase equalizes differential lines
- Energy/bit = $2C(V_{DD}/n)^2$
- Challenges: Receiver design, noise margins

[Ref: H. Yamauchi, JSSC’95]
Alternative Communication Schemes

- Example: Capacitively driven wires

\[
C_c = \frac{(C_{\text{wire}} + C_{\text{load}})}{n}
\]

where \( n = 10-20 \)

- Offers some compelling advantages
  - Reduced swing
    - Swing is \( V_{\text{cc}}/(n+1) \) without extra supply
  - Reduced load
  - Allows for smaller driver
  - Reduced delay
  - Capacitor pre-emphasizes edges

[Ref: D. Hopkins, ISSCC’07]
Signaling Protocols

Globally Asynchronous
self-timed handshaking protocol

Allows individual modules to dynamically trade-off performance for energy efficiency

Slide 6.27
Signaling Protocols

Network

d_req, ack

Physical Layer Interface Module

d_in, d_out
clk done

Processor Module
(mProc, ALU, MPY, SRAM…)

d_in, req

Globally asynchronous

d_in, req
clk done

Locally synchronous

Slide 6.28
The Data Link/Media Access Layer

Reliable transmission over physical link and sharing interconnect medium between multiple sources and destinations (MAC)
- Bundling, serialization, packetizing
- Error detection and correction
- Coding
- Multiple-access schemes
Adding redundancy to communication link (extra bits) to:
- Reduce transitions (activity encoding)
- Reduce energy/bit (error-correcting coding)
Activity Reduction Through Coding

Example: Bus-Invert Coding

- Data word $D$ inverted if Hamming distance from previous word is larger than $N/2$.

<table>
<thead>
<tr>
<th>$D$</th>
<th>$# T$</th>
<th>$D_{enc}$</th>
<th>$p$</th>
<th>$# T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>00101010</td>
<td>-</td>
<td>00101010</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>00111011</td>
<td>2</td>
<td>00111011</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>11010100</td>
<td>7</td>
<td>00101011</td>
<td>1</td>
<td>1+1</td>
</tr>
<tr>
<td>00001101</td>
<td>5</td>
<td>00001101</td>
<td>0</td>
<td>3+1</td>
</tr>
<tr>
<td>01110110</td>
<td>6</td>
<td>10001001</td>
<td>1</td>
<td>2+1</td>
</tr>
</tbody>
</table>

[Ref: M. Stan, TVLSI'95]

Slide 6.31
Bus-Invert Coding

Encode

Bus

Decode

Gain:
- 25% (at best – for random data)

Overhead:
- Extra wire (and activity)
- Encoder, decoder
- Not effective for correlated data

[Ref: M. Stan, TVLSI'95]
Other Transition Coding Schemes

- Advanced bus-invert coding (e.g. partition bus into sub-components) (e.g. [M. Stan, TVLSI’97])
- Coding for address busses (which often display sequentiality) (e.g. [L. Benini, DATE’98])
- Full-fledged channel coding, borrowed from communication links (e.g. [S. Ramprasad, TVLSI’99])

<table>
<thead>
<tr>
<th>bit k-1</th>
<th>bitk</th>
<th>bit k+1</th>
<th>Delay factor $g$</th>
</tr>
</thead>
<tbody>
<tr>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>1</td>
</tr>
<tr>
<td>↑</td>
<td>↑</td>
<td>↓</td>
<td>$1 + r$</td>
</tr>
<tr>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>$1 + 2r$</td>
</tr>
<tr>
<td>↓</td>
<td>↑</td>
<td>↓</td>
<td>$1 + 3r$</td>
</tr>
<tr>
<td>↓</td>
<td>↑</td>
<td>↑</td>
<td>$1 + 4r$</td>
</tr>
</tbody>
</table>

Coding to reduce impact of Miller capacitance between neighboring wires
[Ref: Sotiriadis, ASPDAC’01]

Maximum capacitance transition—can be avoided by coding
Error-Correcting Codes

Example: (4,3,1) Hamming Code

\[
\begin{align*}
P_1 P_2 B_3 P_4 B_5 B_6 B_7 &= 0 \\
B_3 \oplus B_5 \oplus B_6 \oplus B_7 &= 0 \\
P_4 \oplus B_6 \oplus B_5 \oplus B_7 &= 0
\end{align*}
\]

- Adding redundancy allows for more aggressive scaling of signal swings and/or timing
- Simpler codes such as Hamming prove most effective

Slide 6.34
Media Access

- Sharing of physical media over multiple data streams increases capacitance and activity (see Chapter 5), but reduces area.
- Many multi-access schemes known from communications
  - Time domain: Time-Division Multiple Access (TDMA)
  - Frequency domain: narrow band, code division multiplexing
- Buses based on Arbitration-based TDMA most common in today’s ICs.

Slide 6.35
Bus Protocols and Energy

- Some lessons from the communications world:
  - When utilization is low, simple schemes are more effective
  - When traffic is intense, reservation of resources minimizes overhead and latency (collisions, resends)
- Combining the two leads to energy efficiency
- Example: Silicon backplane micro-network

Independent arbitration for every cycle includes two phases:
- Distributed TDMA for guaranteed latency/bandwidth
- Round-robin for random access

[Courtesy: Sonics, Inc]
The Network Layer

Topology-independent end-to-end communication over multiple data links (routing, bridging, repeaters)
- Topology
- Static versus dynamic configuration/routing

Becoming more important in today’s complex multi-processor designs “The Network-on-a-Chip (NoC)”

[Ref: G. De Micheli, Morgan-Kaufman'06]
Network-on-a-Chip (NoC)

- Dedicated networks with reserved links preferable for high-traffic channels – but limited connectivity, area overhead
- Flexibility an increasing requirement in (many) multi-core chip implementations
The Network Trade-offs

Interconnect-oriented architecture trades off flexibility, latency, energy and area efficiency through the following concepts:

- Locality – eliminate *global* structures
- Hierarchy – expose locality in communication requirements
- Concurrency/Multiplexing

*Very Similar to Architectural Space Trade-offs*

[Image: Diagram showing a network-on-a-chip with labels for Local Logic, Router, Network, Wires, Proc, and a note saying “Dedicated wiring”]

[Courtesy: B. Dally, Stanford]
Networking Topologies

- **Homogeneous**
  - Crossbar, Butterfly, Torus, Mesh, Tree, …

- **Heterogeneous**
  - Hierarchy

Slide 6.40
Network Topology Exploration

Short connections in tree are redundant

Inverse clustering complements mesh

[Ref: V. George, Springer'01]
Circuit-Switched Versus Packet-Based

- **On-chip reality**: Wires (bandwidth) are relatively cheap; buffering and routing expensive
- Packet-switched approach versatile
  - Preferred approach in large networks
  - But ... routers come with large overhead
  - Case study Intel: 18% of power in link, 82% in router
- Circuit-switched approach attractive for high-data-rate quasi-static links
- Hierarchical combination often preferred choice

Hierarchical circuit- and packet-switched networks for longer connections

Slide 6.42
Example: The Pleiades Network-on-a-Chip

• Configurable platform for low-energy communication and signal-processing applications (see Chapter 5)
• Allows for dynamic task-level reconfiguration of process networks

Energy-efficient flexible network essential to the concept

[Ref: H. Zhang, JSSC’00]
Pleiades Network Layer

Hierarchical reconfigurable mesh network

- Network statically configured at start of session and ripped up at end
- Structured approach reduces interconnect energy by a factor of 7 over straightforward crossbar
Top Layers of the OSI Stack

- Abstracts communication architecture to system and performs data formatting and conversion
- Establishes and maintains end-to-end communications
  - flow control, message reordering, packet segmentation and reassembly

Example: Establish, maintain, and rip up connections in dynamically reconfigurable systems-on-a-chip – Important in power management

Slide 6.45
What about Clock Distribution?

- Clock easily the most energy-consuming signal of a chip
  - Largest length
  - Largest fan-out
  - Most activity ($\alpha = 1$)

- Skew control adding major overhead
  - Intermediate clock repeaters
  - De-skewing elements

- Opportunities
  - Reduced swing
  - Alternative clock distribution schemes
  - Avoiding a global clock altogether
Reduced-Swing Clock Distribution

- Similar to reduced-swing interconnect
- Relatively easy to implement
- But extra delay in flip-flops adds directly to clock period

Example: half-swing clock distribution scheme

[Ref: H. Kojima, JSSC’95]
Alternative Clock Distribution Schemes

Example: Transmission-Line Based Clock Distribution

Canceling skew in perfect transmission line scenario

[Ref: V. Prodanov, CICC’06]
Summary

- Interconnect important component of overall power dissipation
- Structured approach with exploration at different abstraction layers most effective
- Lot to be learned from communications and networking community – yet, techniques must be applied judiciously
  - Cost relationship between active and passive components different
- Some exciting possibilities for the future: 3D integration, novel interconnect materials, optical or wireless I/O
Books and Book Chapters


Articles

References (cont.)

References (cont.)