Digital Integrated Circuits
A Design Perspective

Jan M. Rabaey
Anantha Chandrakasan
Borivoje Nikolic

Coping with Interconnect

December 15, 2002
Impact of Interconnect Parasitics

- Reduce Robustness
- Affect Performance
  - Increase delay
  - Increase power dissipation

Classes of Parasitics
- Capacitive
- Resistive
- Inductive
INTERCONNECT

Dealing with Capacitance
Capacitive Cross Talk

\[ \Delta V_Y = \frac{C_{XY}}{C_Y + C_{XY}} \Delta V_X \]
Capacitive Cross Talk
Dynamic Node

3 x 1 \( \mu m \) overlap: 0.19 V disturbance

© Digital Integrated Circuits 2nd Interconnect
Capacitive Cross Talk
Driven Node

Keep time-constant smaller than rise time

\[ \tau_{XY} = R_Y(C_{XY} + C_Y) \]

© Digital Integrated Circuits 2nd Interconnect
Dealing with Capacitive Cross Talk

- Avoid floating nodes
- Protect sensitive nodes
- Make rise and fall times as large as possible
- Differential signaling
- Do not run wires together for a long distance
- Use shielding wires
- Use shielding layers
Shielding

Shielding wire

GND

V_{DD}

Shielding layer

Substrate (GND)
Cross Talk and Performance

- When neighboring lines switch in opposite direction of victim line, delay increases

DELAY DEPENDENT UPON ACTIVITY IN NEIGHBORING WIRES

Miller Effect

- Both terminals of capacitor are switched in opposite directions
  \((0 \rightarrow V_{dd}, V_{dd} \rightarrow 0)\)

- Effective voltage is doubled and additional charge is needed
  \((from \ Q=CV)\)
# Impact of Cross Talk on Delay

<table>
<thead>
<tr>
<th>bit $k-1$</th>
<th>bit $k$</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>$1 + 2r$</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>

$r$ is ratio between capacitance to GND and to neighbor.
Structured Predictable Interconnect

Example: Dense Wire Fabric ([Sunil Kathri])

Trade-off:
- Cross-coupling capacitance 40x lower, 2% delay variation
- Increase in area and overall capacitance

Also: FPGAs, VPGAs
Interconnect Projections
Low-k dielectrics

- Both *delay and power are reduced* by dropping interconnect capacitance
- Types of low-k materials include: inorganic (SiO₂), organic (Polyimides) and aerogels (ultra low-k)
- The numbers below are on the conservative side of the NRTS roadmap

<table>
<thead>
<tr>
<th>Generation</th>
<th>0.25 μm</th>
<th>0.18 μm</th>
<th>0.13 μm</th>
<th>0.1 μm</th>
<th>0.07 μm</th>
<th>0.05 μm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dielectric Constant</td>
<td>3.3</td>
<td>2.7</td>
<td>2.3</td>
<td>2.0</td>
<td>1.8</td>
<td>1.5</td>
</tr>
</tbody>
</table>
Encoding Data Avoids Worst-Case Conditions

![Diagram of encoder and decoder with input and output signals](image-url)
Driving Large Capacitances

\[ t_p = \frac{C_L V_{swing}}{I_{av}} \]

- Transistor Sizing
- Cascaded Buffers
Using Cascaded Buffers

0.25 μm process
$Cin = 2.5 \text{ fF}$
$tp0 = 30 \text{ ps}$

$F = \frac{CL}{Cin} = 8000$
$f_{opt} = 3.6 \quad N = 7$
$tp = 0.76 \text{ ns}$

(See Chapter 5)
Output Driver Design

Trade off Performance for Area and Energy

Given $t_{pmax}$ find $N$ and $f$

- **Area**
  
  \[ A_{\text{driver}} = \left(1 + f + f^2 + \ldots + f^{N-1}\right) A_{\min} = \frac{f^N - 1}{f - 1} A_{\min} = \frac{F - 1}{f - 1} A_{\min} \]

- **Energy**
  
  \[ E_{\text{driver}} = \left(1 + f + f^2 + \ldots + f^{N-1}\right) C_i V_{DD}^2 = \frac{F - 1}{f - 1} C_i V_{DD}^2 \approx \frac{C_L}{f - 1} V_{DD}^2 \]
Delay as a Function of $F$ and $N$

![Graph showing delay as a function of $F$ and $N$.]
Output Driver Design

0.25 μm process, $C_L = 20 \ pF$

Transistor Sizes for optimally-sized cascaded buffer $t_p = 0.76 \ ns$

<table>
<thead>
<tr>
<th>Stage</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W_n$ (μm)</td>
<td>0.375</td>
<td>1.35</td>
<td>4.86</td>
<td>17.5</td>
<td>63</td>
<td>226.8</td>
<td>816.5</td>
</tr>
<tr>
<td>$W_p$ (μm)</td>
<td>0.71</td>
<td>2.56</td>
<td>9.2</td>
<td>33.1</td>
<td>119.2</td>
<td>429.3</td>
<td>1545.5</td>
</tr>
</tbody>
</table>

Transistor Sizes of redesigned cascaded buffer $t_p = 1.8 \ ns$

<table>
<thead>
<tr>
<th>Stage</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W_n$ (μm)</td>
<td>0.375</td>
<td>7.5</td>
<td>150</td>
</tr>
<tr>
<td>$W_p$ (μm)</td>
<td>0.71</td>
<td>14.4</td>
<td>284</td>
</tr>
</tbody>
</table>
How to Design Large Transistors

- Reduces diffusion capacitance
- Reduces gate resistance

Small transistors in parallel
Bonding Pad Design

Bonding Pad

100 μm

© Digital Integrated Circuits 2nd Interconnect
ESD Protection

- When a chip is connected to a board, there is unknown (potentially large) static voltage difference.
- Equalizing potentials requires (large) charge flow through the pads.
- Diodes sink this charge into the substrate – need guard rings to pick it up.
ESD Protection

Diode
Chip Packaging

- Bond wires (~25μm) are used to connect the package to the chip
- Pads are arranged in a frame around the chip
- Pads are relatively large (~100μm in 0.25μm technology), with large pitch (100μm)
- Many chips areas are ‘pad limited’
Pad Frame

Layout

Die Photo

© Digital Integrated Circuits 2nd

Interconnect
Chip Packaging

- An alternative is ‘flip-chip’:
  - Pads are distributed around the chip
  - The soldering balls are placed on pads
  - The chip is ‘flipped’ onto the package
  - Can have many more pads
Tristate Buffers

\[ \text{Out} = \text{In}.\text{En} + \bar{\text{En}}.\text{Z} \]

Increased output drive
Reducing the swing potentially yields linear reduction in delay.

Also results in reduction in power dissipation.

Delay penalty is paid by the receiver.

Requires use of “sense amplifier” to restore signal level.

Frequently designed differentially (e.g. LVDS).
Single-Ended Static Driver and Receiver
Dynamic Reduced Swing Network

\[ V_{bus}, V_{sym}, V_{asym} \]

\[ V(Volt) \]

\[ f \]
Dealing with Resistance
Impact of Resistance

- We have already learned how to drive RC interconnect
- Impact of resistance is commonly seen in power supply distribution:
  - IR drop
  - Voltage variations
- Power supply is distributed to minimize the IR drop and the change in current due to switching of gates
RI Introduced Noise
Power Dissipation Trends

- Power consumption is increasing
  - Better cooling technology needed
- Supply current is increasing faster!
- On-chip signal integrity will be a major issue
- Power and current distribution are critical
- Opportunities to slow power growth
  - Accelerate Vdd scaling
  - Low κ dielectrics & thinner (Cu) interconnect
  - SOI circuit innovations
  - Clock system design
  - micro-architecture

ASP DAC 2000
Resistance and the Power Distribution Problem

Before

- Requires fast and accurate peak current prediction
- Heavily influenced by packaging technology

After

© Digital Integrated Circuits2nd
Source: Cadence
Power Distribution

- Low-level distribution is in Metal 1
- Power has to be ‘strapped’ in higher layers of metal.
- The spacing is set by IR drop, electromigration, inductive effects
- Always use multiple contacts on straps
Power and Ground Distribution

(a) Finger-shaped network

(b) Network with multiple supply pins
3 Metal Layer Approach (EV4)

3rd “coarse and thick” metal layer added to the technology for EV4 design

Power supplied from two sides of the die via 3rd metal layer

2nd metal layer used to form power grid

90% of 3rd metal layer used for power/clock routing

© Digital Integrated Circuits 2nd

Courtesy Compaq
4 Metal Layers Approach (EV5)

4th “coarse and thick” metal layer added to the technology for EV5 design
Power supplied from four sides of the die
Grid strapping done all in coarse metal
90% of 3rd and 4th metals used for power/clock routing

© Digital Integrated Circuits2nd

Courtesy Compaq

Interconnect
6 Metal Layer Approach – EV6

2 reference plane metal layers added to the technology for EV6 design
- Solid planes dedicated to Vdd/Vss
- Significantly lowers resistance of grid
- Lowers on-chip inductance

© Digital Integrated Circuits2nd

Courtesy Compaq
Electromigration (1)

Limits dc-current to 1 mA/μm
Electromigration (2)
Resistivity and Performance

The distributed rc-line

Diffused signal propagation

Delay $\sim L^2$
The Global Wire Problem

\[ T_d = 0.377 R_w C_w + 0.693 \left( R_d C_{out} + R_d C_w + R_w C_{out} \right) \]

Challenges

- No further improvements to be expected after the introduction of Copper (superconducting, optical?)
- Design solutions
  - Use of fat wires
  - Insert repeaters — but might become prohibitive (power, area)
  - Efficient chip floorplanning
- Towards “communication-based” design
  - How to deal with latency?
  - Is synchronicity an absolute necessity?
Interconnect Projections: Copper

- Copper is planned in full sub-0.25 μm process flows and large-scale designs (IBM, Motorola, IEDM97)
- With cladding and other effects, Cu ~ 2.2 μΩ-cm vs. 3.5 for Al(Cu) ⇒ 40% reduction in resistance
- Electromigration improvement; 100X longer lifetime (IBM, IEDM97)
  - Electromigration is a limiting factor beyond 0.18 μm if Al is used (HP, IEDM95)
**Interconnect: # of Wiring Layers**

The number of metal layers is steadily increasing due to:

- Increasing die size and device count: we need more wires and longer wires to connect everything.
- Rising need for a hierarchical wiring network; local wires with high density and global wires with low RC.

---

### Minimum Widths (Relative)

<table>
<thead>
<tr>
<th>Layer</th>
<th>Width (µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M5</td>
<td>1.0</td>
</tr>
<tr>
<td>M4</td>
<td>1.25</td>
</tr>
<tr>
<td>M3</td>
<td>1.5</td>
</tr>
<tr>
<td>M2</td>
<td>1.75</td>
</tr>
<tr>
<td>M1</td>
<td>2.0</td>
</tr>
</tbody>
</table>

### Minimum Spacing (Relative)

<table>
<thead>
<tr>
<th>Layer</th>
<th>Spacing (µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M5</td>
<td>0.75</td>
</tr>
<tr>
<td>M4</td>
<td>0.8</td>
</tr>
<tr>
<td>M3</td>
<td>0.9</td>
</tr>
<tr>
<td>M2</td>
<td>1.0</td>
</tr>
<tr>
<td>M1</td>
<td>1.2</td>
</tr>
</tbody>
</table>

---

© Digital Integrated Circuits 2nd
Diagonal Wiring

- 20+% Interconnect length reduction
- Clock speed
  - Signal integrity
  - Power integrity
- 15+% Smaller chips
  - plus 30+% via reduction

© Digital Integrated Circuits 2nd

Courtesy Cadence X-initiative

Interconnect
Using Bypasses

Using a metal bypass

Driving a word line from both sides

Metal word line

Polysilicon word line

Driver

WL

K cells
Reducing RC-delay

Repeater

\[ M = L \sqrt{\frac{0.38rc}{t_{pbuf}}} \]  

(chapter 5)
Repeater Insertion (Revisited)

Taking the repeater loading into account

\[ m_{opt} = L \sqrt{\frac{0.38rc}{0.69R_cC_d(\gamma + 1)}} = \sqrt{\frac{t_{\text{wire(unbuffered)}}}{t_{p1}}} \]

\[ s_{opt} = \sqrt{\frac{R_cC_d}{rC_d}} \]

For a given technology and a given interconnect layer, there exists an optimal length of the wire segments between repeaters. The delay of these wire segments is independent of the routing layer!

\[ L_{crit} = \frac{L}{m_{opt}} = \sqrt{\frac{t_{p1}}{0.38rc}} \quad t_{p, crit} = \frac{t_{p, min}}{m_{opt}} = 2 \left( 1 + \sqrt{\frac{0.69}{0.38(1 + \gamma)}} \right) t_{p1} \]
Dealing with Inductance
Impact of inductance on supply voltages:

- Change in current induces a change in voltage
- Longer supply lines have larger $L$
**L di/dt: Simulation**

![Graphs showing L di/dt simulations with and without inductors, comparing input rise/fall times of 50 psec and 800 psec.](image)

**Without inductors**: The graphs show a smooth and gradual voltage rise and fall with minimal overshoot.

**With inductors**: The presence of inductors introduces more pronounced overshoot and transient behavior, as evident from the more oscillatory voltage response.

**Input rise/fall time**: 50 psec vs. 800 psec.
Dealing with $\frac{dI}{dt}$

- Separate power pins for I/O pads and chip core.
- Multiple power and ground pins.
- Careful selection of the positions of the power and ground pins on the package.
- Increase the rise and fall times of the off-chip signals to the maximum extent allowable.
- Schedule current-consuming transitions.
- Use advanced packaging technologies.
- Add decoupling capacitances on the board.
- Add decoupling capacitances on the chip.
Choosing the Right Pin
Decoupling capacitors are added:
- on the board (right under the supply pins)
- on the chip (under the supply straps, near large buffers)
**De-coupling Capacitor Ratios**

- **EV4**
  - total effective switching capacitance = 12.5nF
  - 128nF of de-coupling capacitance
  - de-coupling/switching capacitance ~ 10x

- **EV5**
  - 13.9nF of switching capacitance
  - 160nF of de-coupling capacitance

- **EV6**
  - 34nF of effective switching capacitance
  - 320nF of de-coupling capacitance -- not enough!

Source: B. Herrick (Compaq)
**EV6 De-coupling Capacitance**

Design for $\Delta I_{dd} = 25 \text{ A} @ V_{dd} = 2.2 \text{ V}$, $f = 600 \text{ MHz}$

- 0.32-$\mu$F of on-chip de-coupling capacitance was added
  - Under major busses and around major gridded clock drivers
  - Occupies 15-20% of die area

- 1-$\mu$F 2-$\text{cm}^2$ Wirebond Attached Chip Capacitor (WACC) significantly increases “Near-Chip” de-coupling
  - 160 Vdd/Vss bondwire pairs on the WACC minimize inductance
389 Signal - 198 VDD/VSS Pins

- 389 Signal Bondwires
- 395 VDD/VSS Bondwires
- 320 VDD/VSS Bondwires

WACC
Microprocessor

Heat Slug

587 IPGA

Source: B. Herrick (Compaq)
The Transmission Line

\[ \frac{2}{\alpha x^2} \frac{\partial^2 v}{\partial x^2} = rc \frac{\partial^2 v}{\partial t^2} + lc \frac{\partial^2 v}{\partial x^2} \]

The Wave Equation
Design Rules of Thumb

- Transmission line effects should be considered when the rise or fall time of the input signal \((t_r, t_f)\) is smaller than the time-of-flight of the transmission line \((t_{flight})\).
  \[
  t_r (t_f) \ll 2.5 \ t_{flight}
  \]

- Transmission line effects should only be considered when the total resistance of the wire is limited:
  \[
  R < 5 \ Z_0
  \]

- The transmission line is considered lossless when the total resistance is substantially smaller than the characteristic impedance,
  \[
  R < \frac{Z_0}{2}
  \]
Should we be worried?

- Transmission line effects cause overshooting and non-monotonic behavior

Clock signals in 400 MHz IBM Microprocessor (measured using e-beam prober) [Restle98]
Matched Termination

Series Source Termination

Parallel Destination Termination

© Digital Integrated Circuits 2nd Interconnect
Segmented Matched Line Driver
Parallel Termination—
Transistors as Resistors

![Parallel Termination Diagram]

- NMOS only
- PMOS only
- NMOS-PMOS
- PMOS with V bias

Normalized Resistance vs. $V_R$ (Volt)

$V_{dd}$

$V_{bb}$

Out

Out
Output Driver with Varying Terminations

Initial design

Revised design with matched driver impedance
The “Network-on-a-Chip”

- Embedded Processors
- Memory Sub-system
- Interconnect Backplane
- Accelators
- Configurable Accelerators
- Peripherals