Putting it all together—

Chip Level Issues
Full Chip Issues

- Noise
- Power Distribution
- Clocking
- Packaging
- Pads
Chip Power Requirements

- Large Scale Chip Power Phenomenal
  - Pentium 4 @ 0.13um has 85A Peak Package Current
  - @ 1.5V requires $.15/85 = 1.8m\Omega$ total power network resistance
  - On-chip peak current risetime is <100pS!
- $I_{DD}$ changes on many time scales (DC to GHz)
Power Coupled Noise

- Droop due to IR drop, Ldi/dt noise and Supply Inductance
- Modulates behavior of Gates
  - Signalling Failure
  - Reduction of Noise Budget (Can you afford dynamic logic)
  - Reduction of System Performance
  - Increase in Power Dissipation
  - Reduction of device reliability
    - Hot Electrons
    - Oxide Damage
    - Electromigration
Noise to Jitter Conversion: Fundamentals

- Uncertainty of threshold reference (A from power supply noise) determines jitter
  - The buffer can switch (threshold) anywhere in this region (A’)
  - The slower the rise time the more opportunity is presented to PWR noise
- Amount of jitter directly proportional to the magnitude of the noise/ripple/GND bounce
  - \( B \text{ (jitter)} = A \text{ (noise) } \times \frac{dt}{dV} \)

© Digital Integrated Circuits 2nd and F. Brewer


Gate Behavior with Noise

- Effective propagation time can be longer or shorter due to noise
  - Delay is proportional to noise magnitude
  - Noise induced delay can be either positive or negative

© Digital Integrated Circuits 2nd and F. Brewer

Interconnect
Logic Current Profile

- Assume triangle current profile: \( Q = C_{Load} V_{dd} \)

- Peak Current
  \[ i_{peak} \approx \frac{2Q}{1.8t_r} = \frac{1.1C_{Load}V_{dd}}{t_r} \]

- Average Current
  \[ i_{avg} = \frac{kC_{Load}V_{dd}}{t_{ui}} \]

- K denotes the probability of switching
  - K=.5 for a clock
  - K=.2 for a heavily used part of microprocessor
  - K=.1 or less for typical ASIC
Power Distribution

- Low-level distribution is in Metal 1
  - Higher Resistance
- 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

© Digital Integrated Circuits 2nd and F. Brewer
Power Distribution Mesh

- Connection point:
- VDD
- VDD pin

Current contribution
Current flowing path

Module A
Module B
Module C
IR Drop

\[ V_{\text{drop}} = I_{\text{peak}} R_{\text{distribution}} \]

- IR drop is proportional to local peak current
  - Peak current reduced by parasitic bypass capacitance
  - Geometry to estimate \( R_{\text{dist}} \)
  - Inductance usually ignored since small compared to IR
    - Capacitive coupling is very large, inductance is the inverse
    - Not true for low resistance busses (e.g. pad frame wiring)

\[ V_{\text{drop}} = \sum_{\text{path}} I_i R_i + L_i \frac{dI_i}{dt} \]

- Local peak strongly affected by synchronization of clocking
  - Intentional skew (DAC ’98 Vittal)
Power Rail IR Drop

\[ V_{\text{drop}}(\text{max}) = \frac{I_{\text{total}} R_{\text{total}}}{8} \]

- Distributed model of current loads and resistance
  - Supply from both sides, assume uniform load
  - Supply from one side, uniform: 4x as large = IR/2

© Digital Integrated Circuits 2nd and F. Brewer
Interconnect
Resistance and the Power Distribution Problem

Before

After

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

Source: Cadence
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 Circuits²nd and F. Brewer

Interconnect
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 Circuits 2nd and F. Brewer
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
Bypass Calculation I

- Essential idea: Local capacitor supplies power for peak to provide lower frequency requirement to next stage of power network
- \( Q = CV = It \) so: \( C = \frac{t*I}{V} \)
- For impulse of total charge \( q \), we have: \( C = \frac{q}{\Delta V} \)
  - E.G. for \( I = 3A \), \( t=1nS \), \( \Delta V=0.1V \) => \( C=30nF \)
  - E.G. for \( q = 120fC \), \( \Delta V=0.1V \) => \( C=1.2pF \)
Parasitic Bypass

- The majority of gates in a circuit do not switch on a given cycle—
  - Others provide low-resistance (few hundred ohms) path from gates (outputs) to one of the supply rails
  - Roughly 40% of total gate capacitance in given area is connected to each supply rail as bypass
  - (0.18um) 20,000 gates/mm², typical gate has 8-12 fF => 200pF/mm² local bypass or 20nF/1cm² die…
Electromigration (1)

Limits dc-current to $1 \text{mA/\mu m}$
Electromigration (2)
Metal Migration

- **Al** \((2.9 \mu \Omega \text{cm M.P. 660 C})\)
  - \(1 \text{mA/}\mu \text{m}^2\) at 60C is average current limit for 10 year MTTF
  - Current density decreases rapidly with temperature

- **Cu** \((1.7 \mu \Omega \text{cm M.P. 1060 C})\)
  - \(10 \text{mA/}\mu \text{m}^2\) at 100C or better (depends on fabrication quality)
  - Density decreases with temperature, but much slower over practical Silicon operation temperatures <120C

- Find Average current through wire – check cross section
  - Be wary of Via’s!! Typical cross-section: 20-40% of minimal wire.
Clocking
Clock Distribution

H-tree

Clock is distributed in a tree-like fashion
More realistic H-tree

[Restle98]
The Grid System

- No rc-matching
- Large power
Example: DEC Alpha 21164

Clock Frequency: 300 MHz - 9.3 Million Transistors

Total Clock Load: 3.75 nF

Power in Clock Distribution network: 20 W (out of 50)

Uses Two Level Clock Distribution:

• Single 6-stage driver at center of chip
• Secondary buffers drive left and right side clock grid in Metal3 and Metal4

Total driver size: 58 cm!
21164 Clocking

- 2 phase single wire clock, distributed globally
- 2 distributed driver channels
  - Reduced RC delay/skew
  - Improved thermal distribution
  - 3.75nF clock load
  - 58 cm final driver width
- Local inverters for latching
- Conditional clocks in caches to reduce power
- More complex race checking
- Device variation
Clock Skew in Alpha Processor
EV6 (Alpha 21264) Clocking
600 MHz – 0.35 micron CMOS

Global clock waveform

- $t_{\text{cycle}} = 1.67$ ns
- $t_{\text{rise}} = 0.35$ ns
- $t_{\text{skew}} = 50$ ps

- 2 Phase, with multiple conditional buffered clocks
  - 2.8 nF clock load
  - 40 cm final driver width
- Local clocks can be gated “off” to save power
- Reduced load/skew
- Reduced thermal issues
- Multiple clocks complicate race checking

© Digital Integrated Circuits 2nd and F. Brewer

Interconnect
21264 Clocking
EV6 Clock Results

GCLK Skew
(at Vdd/2 Crossings)

GCLK Rise Times
(20% to 80% Extrapolated to 0% to 100%)

© Digital Integrated Circuits²nd and F. Brewer

Interconnect
**EV7 Clock Hierarchy**

**Active Skew Management and Multiple Clock Domains**

- Widely dispersed drivers
- DLLs compensate static and low-frequency variation
- Divides design and verification effort
- DLL design and verification is added work
- Tailored clocks

© Digital Integrated Circuits 2nd and F. Brewer
Capacitive Coupling (Coupling Cross Talk)

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

Keep time-constant smaller than rise time

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

Shielding layer

Substrate (GND)

\( V_{DD} \)
Inductive Issues in Clocking

- Transmission line effects cause overshooting and non-monotonic behavior
- Problem tends to scale out with increased wire resistance
- Substantial problems in power and clock grids due to low resistance

Clock signals in 400 MHz IBM Microprocessor (measured using e-beam prober) [Restle98]
Packaging and Board Interface
Packaging Requirements

- Electrical: Low parasitics
- Mechanical: Reliable and robust
- Thermal: Efficient heat removal
- Economical: Cheap
Bonding Techniques

Wire Bonding

- Substrate
- Die
- Pad
- Lead Frame
Tape-Automated Bonding (TAB)

(a) Polymer Tape with imprinted wiring pattern.

(b) Die attachment using solder bumps.
Flip-Chip Bonding

- Die
- Solder bumps
- Interconnect layers
- Substrate
Cu Flip-Chip Technology
Package-to-Board Interconnect

(a) Through-Hole Mounting  
(b) Surface Mount
Package Types
# Package Parameters

<table>
<thead>
<tr>
<th>Package Type</th>
<th>Capacitance (pF)</th>
<th>Inductance (nH)</th>
</tr>
</thead>
<tbody>
<tr>
<td>68 Pin Plastic DIP</td>
<td>4</td>
<td>35</td>
</tr>
<tr>
<td>68 Pin Ceramic DIP</td>
<td>7</td>
<td>20</td>
</tr>
<tr>
<td>256 Pin Pin Grid Array</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>Wire Bond</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Solder Bump</td>
<td>0.5</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Typical Capacitances and Inductances of Various Package and Bonding Styles (from [Sze83])
Multi-Chip Modules
Package Parasitics

- Use many $V_{DD}$, GND in parallel
  - Inductance, $dl/dt$, Impedance Control
Power System Model

- Simulate system for time and frequency responses
- Power comes from regulator on system board
  - Board and package add parasitic R and L
  - Bypass capacitors help stabilize supply voltage
  - But capacitors also have parasitic R and L
Imperfect Bypass Capacitors

- Even with the addition of bypass capacitance there are still sources of inductance in the current loop which can cause power supply noise.
  - Plane inductance
    - Determined by the shape of the plane (pH/sq) and dielectric thickness
    - E.g. 15cm radius to 2cm radius = 7
  - Bypass capacitor parasitics
  - Capacitor Mounting
    - Solder land, trace to via, *via itself*
Chip Bypass Capacitors

- Series Resistance can create alternative breaks:
  - Often need to parallel capacitors to achieve lower inductance

\[ \omega_{RC} = \frac{1}{RC} \quad \omega_{LR} = \frac{R}{L} \quad \omega_{cap} = \frac{1}{\sqrt{LC}} \]
**Power Supply Inductance**

- Average current through inductor subject to low frequency variations
  - Must control excursions of voltage across the capacitor
  - Inductor does not see high frequency components as long as capacitor can supply bulk of current
  - MUST stay away from resonant frequency of LC circuit
Bypass Reprise: LC step response

- Steps in current trigger resonant response
  \[ V_c = \frac{Q_L}{C} \quad V_L = L \frac{dI(t)}{dt} \quad V_c = V_L \implies V_c = LC \frac{d^2V_c}{dt^2} \]

- Solution:
  \[ V_c(t) = \frac{I_{avg}}{\omega_c C} \sin(\omega_c t) + V_{dd} \]

- Solving for C given restriction on V:
  \[ C > L \left( \frac{I_{avg}}{V_{droop}} \right)^2 \]
Basic Bypass Rules

- Use small capacitor packages
  - Parasitic L is proportional to pkg. Size and aspect ratio
- Use largest value subject to resonant point
  - L is dominated by pkg, so choose C at limit of frequency
- Connect cap lands directly to planes
- NEVER share cap vias
- Keep trace between land and via short!!
  - Benefit of small package is lost otherwise
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!
**EV6 De-coupling Capacitance**

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

- 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-cm$^2$ Wirebond Attached Chip Capacitor (WACC) significantly increases “Near-Chip” de-coupling
  - 160 $V_{dd}/V_{ss}$ bondwire pairs on the WACC minimize inductance
PCB Signaling

- Circuit traces on PCB are transmission lines
  - Impedance Matching
  - Resistor model during drive
  - Pad Rise/Fall times are independent of far-end load in many cases
  - Scale: 6-8in/μS
The Transmission Line

\[ \frac{2}{\partial t} \frac{\partial v}{\partial x^2} = r c \frac{\partial^3 v}{\partial t \partial x} + l c \frac{\partial^2 v}{\partial t^2} \]

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,

\[
R < \frac{Z_0}{2}
\]
**Matched Termination**

**Series Source Termination**

**Parallel Destination Termination**
Parallel Termination—Transistors as Resistors

NMOS only
PMOS only
NMOS-PMOS
PMOS with -1V bias

© Digital Integrated Circuits 2nd and F. Brewer
Interconnect
Output Driver with Varying Terminations

© Digital Integrated Circuits 2nd and F. Brewer
Interconnect
Impact of inductance on supply voltages:
• Change in current induces a change in voltage
• Longer supply lines have larger $L$
• Note: idealized model for decoupling Capacitor!
  • Given model is ok for on-chip
  • Ignores off-chip package inductances of both chip and capacitor
• Decoupling solutions need to cover all time scales of switching
  • Board-level (to 10’s of MHz)
  • Package-level (200 MHz)
  • On-Chip (above 200 MHz)
L $\frac{di}{dt}$: Simulation

- Without inductors
- With inductors
- Decoupled

Input rise/fall time: 50 psec

© Digital Integrated Circuits 2nd and F. Brewer

Interconnect
Simultaneous Switching Noise

- Issue: hundreds of I/O pins
  - Each pin drives 30-90Ω trans-line on PCB
- Potential for very large $\frac{dI}{dt}$ spike:
  \[ \frac{dI}{dt} = \frac{V_{\text{swing}}}{50\Omega \times 1.8t_r} \approx \frac{28mA}{t_r} \]
  \[ L < \frac{V_{\text{droop}}}{dI / dt} = t_r \frac{V_{\text{droop}}}{N_{\text{pins}} \times 28mA} \]
- Eg. 333MHz DDR – 80pins at $t_r = 0.5nS$ (50%):
  - 4.5GA/s => at 0.3V drop, need 63pH power supply inductance
Dealing with $L_{di/dt}$

- Separate power pins for I/O pads and chip core.
  - ESD issues!
- Multiple, interleaved, 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.
- Advanced (TM-line) packaging
- Add decoupling capacitances on the board, package and chip.
389 Signal - 198 VDD/VSS Pins

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

WACC Microprocessor

Heat Slug

587 IPGA
Pads
Pads-- Chip to Board Interface

- Pads drive large Capacitances
  - 5pf minimum to much larger
  - Rise time control
- Board Impedance and Noise
  - L dI/dt Noise
- Coupling to Power Distribution
- ESD
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 and F. Brewer

Interconnect
Pad Example

- Multiple busses provide clean/driver power
- VDD/GND pads drive the busses
- Output pads have protection circuitry and driver circuitry
- Input pads have protection circuitry
- Seal Ring
- Guard Rings

© Digital Integrated Circuits 2nd and F. Brewer

Interconnect
Bus Detail

- Multiple supply rings simplify pad design
- Generic Layout
  Simplifies custom tuning
- Guard Rings Between sections of pad
- ESD/Driver
- Controller
Seal Ring

- Seal Ring is essentially a guard ring with metal layers and contacts placed to lower overglass to substrate evenly at chip boundary.
- Hermetic seal of chip from atmosphere and other contamination.
**Pad Frame**

- Large Power Busses
  Surround Die
- ESD in PADS
- Driver/Logic in Pads
- Seal Ring
- Drive Bypass
Chip to Board Interface -- Pad Design

- Buffer to drive PCB-scale parasitics
  - Capacitance 5-50pF, Impedance 30-90Ω
- Rise-Time Control
  - Noise injection to circuits and power supply
- ESD
  - Protection of chip-scale components
- Perimeter Pads/Area Bump
Driving Large Capacitances

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

- Transistor Sizing
- Cascaded Buffers
Using Cascaded Buffers

\[ \text{In} \quad \Box \quad \Box \quad \Box \quad \text{Out} \]

\[ 1 \quad 2 \quad N \]

\[ C_L = 20 \, \text{pF} \]

0.25 \, \mu \text{m process} \quad F = CL/Cin = 8000
\[ Cin = 2.5 \, \text{fF} \quad f_{opt} = 3.6 \quad N = 7 \]
\[ tp0 = 30 \, \text{ps} \quad tp = 0.76 \, \text{ns} \]

(See Chapter 5)
Output Driver Design

Trade off Performance for Area and Energy

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

- **Area**

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

- **Energy**

  \[ E_{\text{driver}} = (1 + f + f^2 + \ldots + f^{N-1})C_iV_{DD}^2 = \frac{F - 1}{f - 1}C_iV_{DD}^2 \approx \frac{C_L}{f - 1}V_{DD}^2 \]
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

Multiple Contacts

Reduces diffusion capacitance
Reduces gate resistance

small transistors in parallel
Bonding Pad Design
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

[Diagram showing ESD protection circuit with components labeled: PAD, R, D1, D2, V_{DD}, X, C, and a diode indicated.]