Modern Physical Design:
Algorithm
Technology
Methodology
(Part IV)

Stan Chow Ammocore
Andrew B. Kahng UCSD
Majid Sarrafzadeh UCLA

Crosstalk Induced Errors

- Transition on an adjoining signal causes unintended logic transition
- Symptom: chip fails (repeatably) on certain logic operations

[Diagram of crosstalk]
Crosstalk Induced Errors

- Timing dependence on crosstalk
  - timing depends on behavior of adjoining signals
  - symptom: timing predictions inaccurate compared to silicon (effect can be large: 3:1 on individual nets)

![Diagram showing crosstalk effect on timing](image)

Effects of Crosstalk: Delay Uncertainty

Thick thresholds:
- non-threshold
- friendly
- unfriendly

Thin and dotted thresholds:
- Series 1
- Series 2
- Series 3
- Series 4
- Series 5
- Series 6
- Series 7
Effects of Crosstalk: Delay Uncertainty

Conventional Flow
Performance Optimization Methodology

- Design Optimization
  - global restructuring optimization -- logic optimization on layout using actual RC, noise peak values etc.
  - localized optimization -- with no structural changes and least layout impact
  - repeater/buffer insertion for global wires

- Physical optimization
  - high fanout net synthesis (e.g., for clock nets); buffer trees to meet delay/skew and fanout requirements
  - automatically determine network topology (# levels, #buffers, and type of buffers)
  - wire sizing, spacing, shielding etc.

- Fixing timing violations automatically
  - fix setup/hold time violations
  - fix maximum slew and fanout violations

Performance Optimization Methodology

- Static methods (when must we go / can we go dynamic?)
- Filtering-Based methods - approximate analyses
- Analysis Best Practices, Macromodels
- Incremental analysis backplane
Performance Optimization Tool Flow

Architectural/Behavioral Design

RTL Design

Floor Planning

Logic Synthesis

Extraction

Placement

Optimization

Global Routing

Detailed Routing

Electrical Analysis/Estimation

EM Analysis

Power Analysis

Timing Analysis

Signal Integrity Analysis

Clock Analysis

Electrical Optimization
Objective: Minimize Congestion and Wire length with timing

Traditional Approaches
- Weight based approaches
- Net1 priority = Net2 Priority

New Approach
- Electrically Correct driven
- Net1 lower priority, longer net
- Net2 higher priority, shorter net

Managing Glitch / Delay Uncertainty

Physical … proximity of the signal
Temporal … noise event occurs timing window
Critical … is the path important
Electrical … driver strength vs pin cap
Managing Complexity

- Use fast models to reduce complexity
  - 1,000,000 net design 1,000 billion pairwise relationships
- Which signals are prone to coupling noise?
- Net isolation techniques
  - electrical
  - critical
  - logical - nodes cannot switch in same or opposite direction
  - temporal
- Quickly focus on the problem nets
- Identify significant aggressors
- Reduce time and space requirements to a manageable size

Information vs. Control in PD Flow

- Floor planning
  - localize major functions of the chip
  - cluster critical nets
- Placement / Optimization
  - find legal location for all cells and pins
  - traditional objective - minimize wire length
  - gate sizing, buffer insertion, pin swap, etc.
- Global routing
  - resolve congestion
  - find best paths for critical nets
- Detail Routing
  - Create the net geometry
Crosstalk Prevention

- Placement phase
  - Don’t know adjacencies, layer assignment, or global route
  - Do know net length, est. wire R/C, driver strength, signal slews
  - Metrics tell if a net is likely to have problems
  - Fixes (size driver, buffer signal) cause no convergence problems

- Global route phase
  - Don’t know adjacencies, only congestion figure
  - Do know layer assignments, better R/C estimates

Performance Validation Integrated Flow

- Logic design too early for signal integrity problems
- Placement/ Global routing best place to achieve timing convergence
- Placement/ Global routing best place to achieve noise immunization
- Detail routing is too late to fix all the signal integrity problems
Glitch/Delay Uncertainty Optimization

- Increase driver strength
  - net’s glitch and delay uncertainty reduced
  - victims of this net could adversely be affected
  - area/power is increased
- Reduce aggressor strength
  - victims of this net would benefit
  - the net becomes more susceptible to noise
  - area/power reduced
- Insert buffer
  - significantly reduces victim’s noise
- Reduce net length

Coupling Impact Analysis

- Accurate estimate of coupling capacitance
  - statistical model:
    - coupling_per_length = f(congestion, layer)
- Efficient measure of impact on noise and delay
  - prune the net-pair space to a manageable size
  - decouple nets for efficient delay computation
  - compute tight upper bounds on noise energy
Wiring Methodologies

- Electrically connect pins on a net
  - 10 million nets per chip
  - 3-4 pins per net
  - 6-7 routing layers
  - Average length of 160 um per net
- Constraints for Router
  - maximum parallel run for the wire
  - minimum spacing
  - nets to be shielded

Early Wiring Estimation

- Routing Prototype Model (RPM)
  - considers routing congestion
  - considers placement density
  - considers wire length and fanouts (net pinouts)
- Pre-route prototype to Post-route within 5%
- Physical locations for all pins
- Global route
  - congestion
  - Layer information
Predictive Models of Net Length

<table>
<thead>
<tr>
<th></th>
<th>Half</th>
<th>Star</th>
<th>Spine</th>
<th>Min</th>
<th>Steiner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perim</td>
<td>O(N)</td>
<td>O(N)</td>
<td>O(N)</td>
<td>O(N^2 logN)</td>
<td>O(N^2 logN)</td>
</tr>
<tr>
<td>Span</td>
<td>Low</td>
<td>Low</td>
<td>Med</td>
<td>Med</td>
<td>High</td>
</tr>
</tbody>
</table>

where N is number of pins connected to the net

Crosstalk Prevention Strategies

- **Placement phase**
  - don’t know adjacencies, layer assignments, or global routes
  - do know net length, est. wire R/C, driver strength, signal slews
  - establish metrics to tell if net is likely to have problems
  - fixes include driver sizing, buffering

- **Global route phase**
  - don’t know adjacencies, but have idea of congestion
  - do know layer assignments, better R/C estimates

- **Can apply timing windows**
  - only consider signals that can change at the same time
  - data comes from static timing analysis

- **Detailed routing - detailed analysis and routing ECOs**

- **N.B.:** In any case, SI brings potential huge infrastructure changes (e.g., statistical centering design w/distributions)
Pre-Route Pruning

- Victim Pruning
  - sensitivity to coupling effects
  - likelihood of being on a critical path (delay analysis only)
- Aggressor Pruning
  - drive strength and load
- Impact Pruning
  - victim-aggressor pairs
  - physical proximity of net pairs
  - relationship in the time domain

Post-Route Pruning

- Geometric/Capacitive coupling
  - depends on magnitude of coupling capacitance as compared to total capacitance of net
  - Typically less than 10% of nets have significant couplings and couplings to only few nets are significant
- Structural Filtering
  - take into consideration the type of gate driving the nets
  - consider direction of signal propagation or location of driver
  - victim net with weak driver is more likely to have glitch
  - aggressor net with strong driver is more likely to cause glitch on victim nets
- Temporal Pruning
  - switching windows between victim and aggressor nets is used
- Functional or Logical Pruning
  - eliminate signals/paths that can never be responsible for noise
- Finally, need detailed analysis after filtering stage
Timing and Logic Dependence for Glitch

- Structured filtering determines the potentially significant couplings
- Need to tie in the timing or switching window information to determine if indeed these are relevant
- Coupled nets switching at non-overlapping intervals
  - coupling capacitance can be converted to a grounded capacitance
- Coupled nets switch simultaneously with overlapped intervals
  - need to analyze both nets with their drivers simultaneously
  - coupling capacitance can be converted to a grounded capacitance with worst-case switch factor

Reducing Conservatism

- Aggressor nets are not uncorrelated
  - timing correlation: some nodes can not switch at the same time
  - noise effect depends on switching windows of nets

- Logic correlation: some nodes can not switch in opposite direction and some nodes switch monotonically in one direction
Integrated Flow for Crosstalk Avoidance

Criteria for Immunizing Against Crosstalk

- Need models for noise estimation and delay uncertainty computation with tight upper bounds
- Need timing windows and logical transition between coupled nets
- Need physical information
  - accurate estimation of RCs, congestion, density, pin location, and layer information in pre-routing stages
- Fidelity of models is important
### Issues for Delay Uncertainty

- Parameters that effect victim net’s gate and interconnect delays
  - aggressor net(s) coupling capacitance
  - fast slew time at aggressor nets and large aggressor net drivers
  - slow slew time on victim net and small victim net drivers
  - proximity of aggressor driver to victim net’s driver (same direction vs opposite direction signal propagation)
  - power bounce in victim driver/receiver gates
- Impact the following delay components
  - victim net driver gate delay
  - victim net delay
  - victim net receiver gate delay
Impact of Coupling on Delay

- Coupling adds delay ambiguity/uncertainty
  - amount of coupling depends on the temporal relationship between waveforms; coupling varies between nodes of distributed coupled RC networks

- Current timing tools employ a technique which convert coupling caps. to grounded caps by multiplying with a switch factor depending upon switching conditions
  - reduces coupled RC networks to just uncoupled RC networks
  - allows efficient interconnect delay analysis
  - introduces error in delay computation
  - can it be applicable for worst-case analysis?

Timing Windows for Crosstalk

- Only consider signals that can change at the same time
- Data comes from static timing analysis

One clock cycle

A

B

C

D

STA Timing Windows

SI Crosstalk Magnitudes

Worst case occurs here, does not include signals A or D.
Crosstalk Analysis Approaches

- Simple worst case model
  - assume all aggressor nets are switching simultaneously in opposite direction to victim net
  - functional and timing information not considered
  - yields pessimistic results

- Static Analysis Model
  - timing information is used but function information is ignored
  - yields accurate results but might flag too many nets as problematic

Zero-Delay Model
- temporal information is ignored
- search functional space to find input vectors causing maximum noise
- if both aggressor and victim nets switch in the same cycle then assumes conservatively that transitions are correlated

Delay-Vector Model
- consider both timing and functional information completely
- not always possible to obtain exact vectors causing maximum noise
Switch Factor Methodology

- Switch factor for coupling cap. between pair of nets:
  - 0 for nets switching in same direction (in phase)
  - 1 when aggressor net is quiet and only victim net is switching
  - 2 for nets switching in opposite direction (out of phase)
- effective change in voltage, $\Delta V_{\text{eff}} = V_A - V_B = 2V_{dd}$

- Switch factor depends on slew times, relative offset of waveforms at both nodes: need accurate models, iterative methods for accurate SF

Standard Practice

- Use SF=2.0 for most coupling nets
  - coupled nets nets which are part of same logical bus
  - coupled clock nets and regular nets coupling to clock nets
  - pessimistic for some nets and could be optimistic?
  - due to disparity between rise/fall slew times switch factor could be more than 2!
Effect of Slew Time on Switch Factor

- SF is in the range (0, 3)

\[ T_v = N \times T_a \]

Parameters Affecting Switch Factor

- All possible combinations of switch factors
  \[ SF_{ij} = f(A_j, V_l, T_{ij}, X_{ij}) \]
  - \( A_j \) represents event times for aggressor nets
  - \( V_l \) represents event times for victim net
  - \( T_{ij} \) is ratio of aggressor and victim slew times
  - \( X_{ij} \) is cross-coupling capacitances between aggressor and victim nets
- SF’s depend on events and transition times and vice-versa
  - Hence, requires iterative solution to find accurate SF’s
IR Drop

- Voltage drop in supply lines from currents drawn by cells
- Symptom: chip malfunctions on certain vectors
- Biggest problem: what’s the worst-case vector?

Voltages depend on currents of other cells

Power supply network consists of wires of varying sizes; they must be big enough, but too big wastes area

IR Drop

- Analysis
  - model I/O P/G supply; C extraction must distinguish decoupling cap between P/G and coupling cap between signals, P/G
- Prevention (good design)
  - P/G lines on same layer, close to each other; large decoupling on chip; process solutions (e.g., DEC Alpha)
Electromigration

- Power supply lines fail due to excessive current
- Symptom: chip eventually fails in the field when wire breaks

Currents depend on driver type, loads, and how often cell is switched
Currents depend on currents of other cells
Current limit depends on wire size

Power supply network consists of wires of varying sizes; they must be big enough, but too big wastes area

Electromigration

- Prevention: wire cross-section to current rules
- Maximum current density for particular material (via, layer)
- Modified Black’s equation; waveform models
- Higher limits for short, thin wires due to grain effects
- Copper: 100x resistance to EM → not a problem any more? (actually, 4-5x)
Hot Electron Effects

- May also be called short channel effect
- Caused by extremely high electric fields in the channel
  - Occurs when voltages are not scaled as fast as dimensions
- Effect becomes worse as devices are turned on harder
- Symptom: Thresholds shift over time until chip fails

Electrons pick up speed in channel; ‘hot’ electrons are the fastest of a statistically fast bunch

Impact ionization occurs here

Hot Electron Effect (cont)

- Depends on how hard device is driven (input slew rate)
- And on the size of the load
Hot Electron Prevention Strategies

- Allowable region for input slew and output load
- Fluence per transition is function of input slew, output load
- Set maximum allowed degradation over life of device
  (estimate of total number of transitions) $\equiv$ fluence limit
- Size device as needed
- Output load vs. driver sizes

SI Flow - Hot Electron

Problems fixed at each step

Crosstalk Prevention
Sig Wire Self Heat
Hot Electron
Clik Wire Self Heat

Sig Wire/ Clik Wire Self Heat
Electro-migration
IR Drop
Crosstalk Parasitics
Crosstalk Delay

System Level Constraints

Verilog VHDL
STA Place PBS CTS Route RCX Delay Calc

Activities

Timing Windows

New Library Data
**SI Flow - Hot Electron**

Problems fixed at each step

- Crosstalk Prevention
- Sig Wire
- Self Heat
- Hot Electron

System Level Constraints

- Verilog
- VHDL
- STA
- Place PBS
- CTS
- Route
- RCX
- Delay Calc

Timing Windows

New Library Data

**New Design Rule for Hot Electron Effects**

- Allowable region for input slew and output load

- Input rise time too long; causes too much short circuit current
- Output rise time too long
- Output load too big
- Hot electron degradation

Legal operating area

Load Capacitance

- Limit may depend on operating frequency
Hot Electron Characterization

- Get raw process degradation
- Select cell EOL performance (5% timing loss over life)
- Use BTA BERT model to compute total acceptable fluence
- Compute fluence per transition as a function (load, slew)
- Compute Total fluence

Total fluence = fluence/transition * freq * lifetime

- Replace Driver, as needed

Wire Self-Heat

- May also be called signal wire electromigration
- Wire heats above oxide temperature as pulses go through
- Symptom: chip eventually fails when wire breaks
- Depends on metal composition, signal frequency, wire sizes, slew rates, and amount of capacitance driven
- Requires different data/formulas from power supply EM