Interconnect Parasitic Extraction in the Digital IC Design Methodology

Mattan Kamon
Microcosm Technologies, Inc
Cambridge, Ma
matt@memcad.com

Steve McCormick
Sapphire Design Automation
Santa Clara, CA
steve@sdai.com

Ken Shepard
Columbia University
New York, NY
shepard@ee.columbia.edu

Abstract

Accurate interconnect analysis has become essential not only for post-layout verification but also for synthesis. This tutorial explores interconnect analysis and extraction methodology on three levels: coarse extraction to guide synthesis, detailed extraction for full-chip analysis, and full 3D analysis for critical nets. We will also describe the electrical issues caused by parasitics and how they have, and will be, influenced by changing technology. The importance of model order reduction will be described as well as methodologies at the synthesis stage for avoiding parasitic problems.

1 Introduction

With increasing circuit speeds and chip densities, consideration of interconnect parasitics must be both more accurate and sophisticated, as well as occur increasingly early in the IC design process. In the design planning stage, this means more accurate estimation techniques to guide design decisions and to enable more "correct-by-construction" synthesis. In the post-layout verification stage, this requires the solution of a complex parasitic extraction and analysis problem on an exploding amount of data. In this paper, we discuss post-layout parasitic extraction as well as approaches to parasitic-aware synthesis.

We begin in Section 2 by discussing the electrical issues that require detailed understanding of parasitics. These include accurate timing, noise, electromigration, and power-supply integrity analysis. We will consider the relative importance of resistance, capacitance, and inductance in these analyses, and also the influence of new technology (e.g. Cu) and scaling. In Section 3, we briefly consider modern approaches to capacitance extraction that include scan-band algorithms combined with analytical or table-look-up models and then consider extending these techniques for inductance.

We briefly enumerate the current approaches to model-order reduction, which has become the essential link between the large number of extracted RLC parasitic circuit elements and practical electrical analysis.

Any RLC extraction method must solve Maxwell's equations at some level. While previously, the need for solution could be limited to the calibration of table-look-up extractors, more recently, increasing frequencies have moved board-level issues on-chip and the lines between circuit theory and field theory have become increasingly blurred. In Section 4, we will discuss efficient approaches to the necessary 3D electromagnetic analysis, from quasistatic to full-wave. Because the extraction can no longer practically be separated from model reduction, we will discuss approaches to generate low order models *directly* from Maxwell's equations.

If a large number of electrical problems are not discovered until after layout, significant redesign and rework will be necessary. As a result, electrical correctness must be engineered into the entire IC design process. In the final sections of this tutorial, we consider how correctness can be built into the interconnect design environment by controlling lengths (through placement), coupling (through congestion control in placement) and circuit immunity to noise (through driver sizing and buffer insertion). Detailed routers must also understand length, width and adjacency criticality. These needs demand accurate interconnect estimates to drive placement, synthesis, and routings, which we also discuss.

2 Electrical verification

All of the important electrical issues that must be considered in the design of deep submicron integrated circuits require detailed understanding of interconnect parasitics. In this section, we consider the requirements that each of these analyses put on interconnect extraction. Specifically, we consider timing, signal integrity, electromigration, and power-supply integrity analysis. In all cases, these analyses are becoming more complex (with a concommitant increase in the accuracy demands on extraction) as a result of technology scaling, both shrinking feature sizes and increasing clock frequencies. Despite the growing importance of interconnect, one must be wary not to forget the silicon devices in the quest for an "interconnect-centric" design flow.

2.1 Timing and signal integrity analysis

Interconnect analysis as it applies to timing and signal integrity seeks to answer three questions, all potentially in the presence of coupled switching activity on other nets:

- What is the effective loading on each driver? This is necessary to optimize the size of the drivers to minimize delay and slew, and in cases in which inductance is important, to prevent the net from being overdriven, which can produce a ringing response.

- What is the delay and slew at the receivers? In the presence of coupling interactions or inductance, the voltage response can be nonmonotonic.

0-7803-5832-X /99/$10.00 ©1999 IEEE.

- What noise will be coupled at the receiver of an otherwise static net due to switching on other nets?

For dimensions above $0.5\mu m$ or until the late 1980's, interconnect densities are such that capacitive coupling is not a significant concern. In addition, wires are wide enough that resistive effects in the interconnect are important on only the longest nets. In static timing analysis, lumped capacitances are used to model the loads. RC delays, where significant, are modelled using Elmore delays[1], [2]. Net delays are added to the path delay. Slews are degraded across the interconnect by a multiplicative factor of the net delay.

As RC delays become more important in the late 1980's and early 1990's with the scaling of interconnect widths, the Elmore-delay interconnect models with lumped capacitance load models are no longer accurate enough for delay analysis. Interconnect densities, however, are still sufficiently low that coupling is not a concern. As a result, any coupling capacitances are treated as equivalent capacitances tied to ground. At about this time, asymptotic waveform evaluation (AWE)[3], particularly as manifest in the RICE package[4], introduces the CAD community to Padé approximants[5]. For timing analysis, this analysis results in the pi-model, pole-residue interconnect macromodel, in which the load is characterized by a pi model (two grounded capacitors connected by a resistor) and the transfer function from the driver to each of the receivers is characterized by several pole-residue pairs. The Standard Parasitic Exchange Format (SPEF), recently approved as an IEEE standard, contains a representation of this model as its "reduced" form. Because driver characterizations are typically still based on a lumped-C load model, an effective-C (C_{eff})[6] becomes common to leverage existing characterizations even with more complex load models[7].

Just as the standards are crystallizing, the growing importance of interconnect capacitive coupling (due to technology scaling) adds further complexity to the landscape. Below $0.25\mu m$, one must consider the effect of coupling capacitance on signal integrity and timing. Several commercial vendors have extraction tools that can now extract accurate distributed coupled RC networks with "3D" consideration of capacitance. Some of the techniques used in these extractors are discussed in Section 3. Managing all the coupling and producing appropriate interconnect macromodels is usually performed in the analysis tools.

We consider one possible approach to managing coupling RC models in timing and noise analysis, which is similar to that of Reference [8]. A *net complex* is created for each nets in the design. The *primary net* of the complex is the net on which we are trying to analyze the timing or noise. The complex also includes *secondary nets* of significant coupling to the primary net. Other nets that couple to the primary net but which are not deemed significant can be lumped together as a single aggressor port, the *remainder*, which can be discarded or scaled as appropriate. In general, these net complexes can be regarded as multi-input, multi-output (MIMO) systems with the drivers acting as ports. If the receivers can be treated an linear capacitors, they can be incorporated into the complex and the receivers monitored as "taps" (that is, nodes on which to monitor the voltage) rather than ports. Otherwise, the complexes can retain the receivers as ports and can accommodate any linear or nonlinear load. Implicit model-order reduction algorithms exist to reduce these complexes to multiport macromodels, including block versions of Lanczos[9] and Arnoldi[10]. Another variant of block Arnoldi reduction, PRIMA[11], guarantees passive macromodels.

These interconnect macromodels are an important part of static noise analysis[12], in determining glitches that could produce functional failures or in identifying excursions above supply or below ground which could overstress devices (and produce reliability problems) or inject minority carriers into the substrate through forward-biased source-drain junctions. These models are also an important part of static timing analysis where simultaneous switching aggressors can influence the effective loading and net delay[13], [14] on the primary net.

With all the focus on the interconnect, it is easy to forget the silicon. In noise analysis, in addition to considering interconnect coupling, it is important to consider noise that can be injected by "coupling" in the devices, in particular charge-sharing noise and noise injected across source-gate or drain-gate capacitances. In timing, while the effects of interconnect coupling on delay are important, there are other "couplings" in the devices which can be just as important and are generally ignored in static timing analysis. In fact, simultaneously-switching gate inputs can increase or decrease the delay significantly over the single-input-changing case. In addition, the switched state (and in some cases recent switching history) of receiving circuits can significantly influence the load capacitance that they represent to driving gates.

As on-chip frequencies continue to increase and copper interconnects[15] gain prominence, inductance is becoming an important new consideration [16]–[18] in the electrical modelling of interconnect for timing and signal-integrity analysis. (The on-chip inductance of the power grid is also important to power-supply integrity analysis, discussed in Section 2.2.) While inductance has long been considered in high-speed board and package designs, the highly-controlled design style (with its generous utilization of power and ground planes) greatly simplifies the analysis. More complex extraction is required to model inductance within the context of the gridded power and ground distribution common to VLSI chips[19]. This will be considered more in Section 3.

2.2 Power supply integrity and electromigration analysis

The design of the power grid can no longer be taken for granted in deep submicron digital ICs. Because of large average and transient current demands, power-supply integrity and electromigration analysis is a standard part of deep submicron design methodologies, although no really good tools or techniques have been developed to accurately model the current demands of tens of millions of transistors. The power grid is usually extracted independently from the signal lines for power supply and electromigration analysis[7], [20]. "Power points" of some sort are traditional defined in the methodology, pins within the power and ground network that separate the "local" power-ground distribution from the global one. "Peak" and "average" currents are

calculated using dynamic (simulation-based) or static (usually probabilistic) approximations and applied at the power points. IR drop and EM analysis is performed with a dc solve (only need to consider the resistances) of the power grid using these current values. Electromigration analysis further requires that branch currents refer back to extracted geometries so that peak, average, and rms current *densities* can be compared again reliability targets. (The move to Cu eases electromigration concerns by offering about $100\times$ better EM reliability at the same current densities as Al(Cu) interconnect.)

Delta-I (or Li/dt) noise on the power and ground lines is of growing concern because di/dt continues to increase generally with technology scaling (despite scaling supply voltage), requiring that the chip and package be able to supply charge quickly from the supply and to sink charge quickly to ground. This charge must come from a hierarchy of decoupling capacitance, where each level supplies as much charge as it can up to its maximum response frequency, at which point the next level up is forced to respond. If the decoupling at a given level is not adequate to deliver the charge fast enough, then that charge may be required too fast from the next level, resulting in a voltage drop across the associated inductance. As a result, explicit thin-oxide decoupling capacitors have become a mainstay in high-frequency design to provide enough charge at the fastest time constants. Because of the high potential for yield problems (shorts) in large-area thin-oxide capacitors, they are usually equipped with a “shut-off” mechanism.

Effective delta-I noise analysis requires extracting the inductance of the power grid (along with its resistance and capacitance) as well as reasonable estimates of the capacitances contributing to on-chip decoupling. The package model must also be included in the analysis. Resonances in the power distribution close to the clock frequency are usually eliminated (by increasing them and ensuring they are adequately damped) by making sure the inductance to the package is low enough (by utilizing C4 technology and/or increasing the number of leads assigned to the power rails). There are typically, however, low frequency resonances (at 50 MHz, for example, for on-chip clock frequencies of 500 MHz) that are not easily eliminated. Noise margins must consider these power-supply ripples. In particular, dynamic nodes should have keeper devices to ensure that they track the power supply.

3 Techniques for capacitance and resistance extraction

Resistance and capacitance extractors in use today which are capable of full-chip extraction all have a shapes processing engine combined with 3D analytical models for capacitance[21]. Since the number of layout shapes N is very large, scanline algorithms [22]–[25] are widely employed since they require on average only $\sqrt{N}$ states in memory. Scanline techniques are well-established for design rule checking, layout-versus-schematic (connectivity) analysis, as well as parasitic extraction[26]. Shapes processing in extraction involves collecting groups of shapes (an *interaction region*) with significant electrostatic (for capacitance) or magnetostatic (for inductance) coupling (i. e., deciding which coupling to keep and which to discard).

Capacitance has the property of being local (that is, electric field lines emanating from one conductor want to terminate on the closest neighbors), one can process capacitance using a simple interaction region that contains an overlapping set of central conductors and the distances to the nearest neighboring conductors on all layers. To handle coupling in the direction perpendicular to the scanline, a scanband is actually created with the width of the scanband chosen to be the largest coupling distance. After the central conductors of an interaction region fall out of the scanband during scanline traversal of the design, the capacitances of the associated interaction region can be calculated. Capacitance extraction can be performed by matching the interactions to precalculated analytic models for common geometries (often parameterized with distance or width)[21]. When an available pattern is not found, Laplace’s equation can be solved for the interaction region using the techniques described in Section 4. Resistance extraction can either precede or follow capacitance extraction.

Scan-band techniques can also be extended to consider inductance extraction. Unlike capacitance, inductance has the potential to be very nonlocal with complex frequency-dependent return paths. By necessity, interaction regions for inductance extraction will be larger than for capacitance extraction. Reference [17] provides one possible set of geometry-based decomposition rules amenable to scanband shapes processing in which the power and ground distribution of the chip is used to define the interaction region.

4 Full 3D extraction

The extractors of the previous section use precalculated solutions for quick lookup during the the extraction process. To calibrate such engines as well as verify their results for critical nets we require fully general 3D extractors or “field solvers”. This section discusses an approach to fast, general, 3D inductance and capacitance extraction in addition to how such approaches can be combined to directly extract reduced order models without the traditional extraction step. Also discussed are the difficulties in extraction as skin effect, distributed effects, and full-wave effects appear on-chip.

4.1 Fast 3D inductance and capacitance solvers

The heuristics necessary to make full-chip inductance and capacitance extraction practical must be developed from an accurate solution of the governing Maxwell’s equations. For this reason, it is necessary to have full 3D field solvers to analyze the complicated geometries of interconnect accurately, yet fast enough to handle a section of the interconnect significantly larger than the “interaction region” on which the heuristic methods are based.

For capacitance extraction, it is common to apply a boundary element approach which requires discretization of only the surfaces of conductors[27], [28]. The goal of capacitance extraction is to extract the capacitance matrix, C, which relates conductor

voltages to charge via $CV = Q$ where $C \in \mathbb{R}^{k \times k}$ for the k condutor problem. Since the charge density over the conductor is rarely uniform, computing this relation requires *solving* the integral form of Laplace's equation k times. If the surface is divided into n charge panels, each with a constant, unknown charge, then solving Laplace's equation involves computing a matrix solution to

$$Pq = v \tag{1}$$

where $q \in \mathbb{R}^n$ are the unknown charges on each panel, $v \in \mathbb{R}^n$ are the known panel voltages, and P_{ij} can be directly computed from the integral relation.

Since P is a dense matrix, it must be solved with iterative solution techniques accelerated by "fast-multipole" or other matrix sparsification algorithms to handle the 100,000+ number of panels necessary for complicated 3D structures (see [29] for many references). This combined approach reduces the computational (and storage) cost of using boundary and volume-element methods to nearly $O(nk)$, where k is the number of conductors.

Much of the full chip extraction technology for capacitance centers around deriving approximate, yet sophisticated parameterizations of capacitance based on many full 3D solutions. The positive side is that the local nature of capacitance, as described in Section 3, allows one to readily discard many of the distant coupling capacitances, and the diagonal-dominance of the C matrix keeps this truncation from destroying the passivity of the equivalent circuit.

The inductance problem is the opposite. Approximations for the inductance are easy, but inductance is not local and the required size of the interaction region comes only after understanding the entire path of current flow.

To explain why approximate extraction is easy, inductance can be extracted as a non-loop quantity using the concept of partial inductance[28]. Assume we wish to extract self and mutual partial inductance values for two "lumps" of an RLC circuit. We wish to relate current to voltage via the $L, R \in \mathbb{R}^{2 \times 2}$ matrices as

$$(R + j\omega L)I = V$$

where I, V are the currents through each conductor and the voltage across, and ω is the radian frequency. Unlike the capacitance problem, the current density and direction is *constant* in each conductor when the frequency is low enough to ignore skin effect. Then the inductance matrix can be computed directly from the integral relation *without* any matrix solution through

$$L_{ij} = \frac{\mu}{4\pi a_i a_j} \int_{V_i} \int_{V_j'} \frac{\mathbf{l_i} \cdot \mathbf{l_j}}{|r - r'|} dV' \, dV \tag{2}$$

where the integral is over the volume of the conductors, r is the position in a given filament, and $\mathbf{l_i}$ is the unit vector in the direction of current flow for conductor i.

At current clock speeds and on-chip geometries, the constant current density is an acceptable assumption for the interconnect wires. The difficulty instead comes in deciding which entries, L_{ij}, to compute and how to stably account for those neglected.

As clock speeds increase, the current density is no longer constant due to skin and proximity effects. Additionally, the current in structures such as the substrate can play a larger role. For non-wire-like structures such as the substrate, neither the current density distribution or the current direction are known *a priori*.

To model these effects accurately now requires matrix solution. The interior of conductors is divided into bundles of filaments, each which carries a constant current density. Computing the equivalent self and mutual inductance between multiple "lumps" now involves first computing the full partial inductance matrix, L_b, between the n filaments of the problem using (2), and then computing a "circuit solution" to enforce current conservation at the junctions between sections of wire lumps [28]. Since L_b is large and dense, practical extraction again requires efficient sparsification approaches for this circuit solve [30].

4.2 Coupled RLC Model Generation

As frequencies increase, and transmission-line effects move on-chip, the number of "rungs" in the fully coupled RLC ladder network must rise to accurately capture the distributed nature of the interconnect. To verify critical nets with accurate 3D extraction requires a two-fold burden: the extractors of Section 4.1 must solve for more capacitive and inductive elements, and model order reduction (MOR) algorithms must reduce larger networks. For this reason the most efficient approach for accurate low order model generation is to skip the extraction step and go directly from Maxwell's equations to a reduced order model.

The approach requires applying model reduction to a system which can be generated directly using the partial inductance L_b and the potential coefficients P instead of extracted quantities. Panels (not capacitors!) are added to nodes of the equivalent circuit described in the previous section for inductance extraction. By then using a nodal analysis for the panel charges, and mesh analysis for the inductances, one can write a mixed mesh-nodal formulation

$$s \begin{bmatrix} L_m & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} I_m \\ V_p \end{bmatrix} = \begin{bmatrix} -R_m & -MA_p \\ P(MA_p)^T & 0 \end{bmatrix} \begin{bmatrix} I_m \\ V_p \end{bmatrix} + \begin{bmatrix} V_s \\ 0 \end{bmatrix} \tag{3}$$

where I_m and V_p are the mesh currents and nodal voltages, M and A_p are the sparse mesh and nodal incidence matrices, $L_m = ML_bM^t$, and $R_m = MRM^t$ where R is the diagonal resistance matrix.

The matrix structure is similar in structure to the Modified Nodal Analysis (MNA) matrices and can be used with PRIMA[11] to directly generate passive reduced order models[31], [32]. The greatest savings is that the matrix-vector products needed to compute the PRIMA models need only compute matrix-vector products with L_b and a small number of matrix solves with P. Thus, we can still use the sparsification techniques described above. For example, consider extracting a 5 port reduced order model that matches 10 moments for a section of interconnect containg 1000 lumps and assume skin effect is negligible. If one extracted first, and then reduced, accurate computation of the capacitance would require 1000 matrix solves involving P. The number of solves by using PRIMA on an appropriate form of (3) is only 5x10 = 50.

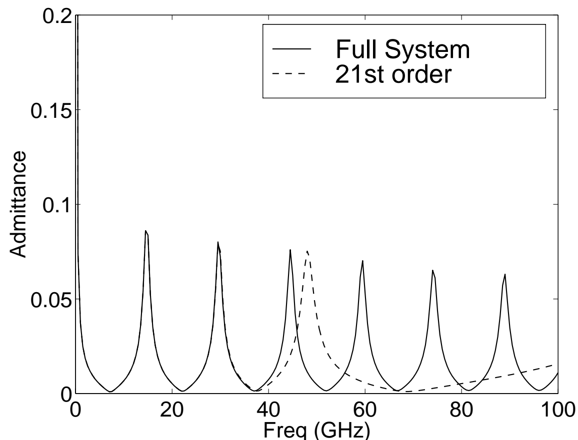

Figure 1: Response of a 1 cm transmission line when modeling skin effect

4.3 Difficulties

Some difficulties arise in the above model reduction process when skin effect is modeled. Consider extracting a reduced model for a 1 cm line transmission line discretized to capture skin effect. The familiar repeated resonance pattern of a transmission line is shown in Figure 1 by computing the frequency response of the original system in (3). Note the decay in the resonant peaks due to skin effect. For this simple response, matching the first two resonances required a 21st order model. One might expect a roughly 4th order model (one conjugate pair per resonance) to be adequate. The difficulty can be attributed to the many "fast modes" associated with the many filaments used to capture skin effect. These poles have small residues, but are very close to the expansion point $s = 0$ so the model reduction approach overemphasizes them. Multipoint approximations are a remedy but are not adequately automatic in point or order selection [32].

Various authors have pursued approaches to avoid the difficulties of skin effect in either the model reduction step or the large number of internal filaments. In [33], a model reduction approach based on an approximation to the optimal Hankel Norm approach is proposed. To avoid generating the many internal filaments for skin effect, in [34], volume elements are replaced with surface elements given the known distribution of the current due to skin and proximity effects. Such an approach has utility when the currents are known to flow parallel to the surfaces. More recently, in [35], a general surface formulation is presented which accounts for *any* current distribution in the volume through only unknowns on the surface. Conductors for which current can flow *in*, such as a substrates, are accurately modelled and are general solutions to Maxwell's equations.

4.4 Full-Wave

For signal integrity issues, coupling has often been localized enough that the propagation time of electromagnetic waves can be assumed to be zero. As circuit speeds increase, the retardation effects may become important for not only signal integrity on chip, but also Electromagnetic Interference (EMI). Adding retardation to the above model reduction approach is difficult because

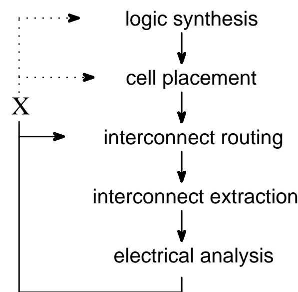

Figure 2: Typical standard-cell design process

P and L_b are now functions of frequency and applying PRIMA directly is no longer possible. Approaches for dealing with such cases are described in [36], [37].

5 Standard Cell Design Process

In this section we will discuss the standard-cell design process and how it relates to design issues that require extraction and electrical analysis, such as timing, coupling noise, and power consumption.

The typical design progresses through the steps outlined in Figure 2. Each of the five steps from logic synthesis through circuit/interconnect analysis is an automated task. Analysis locates the interconnect problems discussed in earlier sections. If the problems detected by analysis occur in manageable numbers, fixes are made, often manually, using the *ECO* (Engineering Change Order) functionality of the design tools.

However, since interconnects are becoming an increasingly significant factor, the number of interconnect related errors is growing to an unmanageable size. Increasingly, a high density of design problems eliminates ECO fixes as an option. As the error count grows, the feasibility of fixing a fully routed design decreases because of the nature of synthesis and placement algorithms. Small input changes lead to significantly different results—while the original errors might be fixed, new ones show up.

5.1 Improvements to Design Process

Newer *electrically directed* algorithms brake the large loop into smaller loops, as shown in Figure 3. Most potential problems can be solved before routing is done, that is, before the actual interconnect and detailed extraction data is known. With electrically directed algorithms, interconnect estimation and electrical analysis *direct* the circuit design, logic design and cell placement algorithms.

The lower loop still exists to fix any residual errors, the expectation being that only a handful of errors will first show up after the final circuit/interconnect analysis.

6 Immunizing a Design from Interconnect Problems

This section considers some operations that electrically directed design tools can take to safeguard a design from interconnect problems before the interconnect is actually designed or routed. Specifically a design tool can:

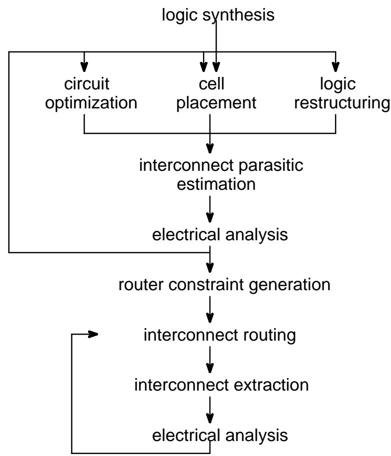

Figure 3: Electrically driven design process

- Keep wire lengths short by controlling placement. The placement algorithm tries to keep any cell close to other cells it connects to; extrapolating onward, it tries to keep every cell close to every other cell. The placement algorithm must always make tradeoff decisions. Traditional placers base the decisions on net connectivity only; much improvement is possible if the decision is also based on timing[38], power and noise criticality.

- Tune driver size for given interconnect. For standard-cell design, this circuit *optimization* process may choose among different cells with similar functionality, or it may insert a buffer with increased drive strength. The optimizer increases driver strengths to reduce delays to the next stage, to sharpen edge times (for power reduction), or to reduce noise glitches generated from switching on coupled nets. It decreases strengths to reduce loading on the previous stage (for either power or timing reasons).

- Break long interconnect spans by inserting buffers[39]. This reduces the quadratic effect of RC interconnect delays. For high fanout nets, multi-level buffer trees may be inserted.

- Structure the logic to better match timing requirements. A design can afford more logic levels along non-critical paths and fewer logic levels along timing critical paths. Restructuring the combinational logic through timing directed boolean minimization[40] may be able to make these tradeoffs and save a few levels of logic on critical paths. This type of analysis and logic structuring is standard with logic synthesis tools.

- Determine physical constraints to pass to the router. It's possible to give modern routers instructions on routing priority, maximum/minimum interconnect lengths, and proximity limits to other nets. Calculating such constraints and passing them to the router may produce a design with more immunity to coupling noise problems.

Several schemes exist for integrating the algorithms in the upper loop of Figure 3. The loop may be constructed of individual programs (e.g., placer, analysis tool, buffer insertion tool)

Action	Possible savings	Impact on other paths
logic restructuring	15 %	high
placement	55 %	high
driver sizing/buffering	30 %	medium
buffer insertion for long interconnects	10 %	low
routing directives	5 %	medium

Table 1: Expected maximum savings for different optimizations

that operate on the entire design in each iteration. Alternatively, the loop may be constructed by tightly integrating the algorithms into one program where individual design decisions are made through one "iteration" of the loop—in this scheme the electrical analysis more closely "directs" the placement, or circuit design. Yet another scheme sets up a large optimization problem that includes a large number of decisions into one set of solvable equations.

Detailed extraction is possible only in the final loop, when the full routing information is known. In the earlier steps, those occurring before interconnect routing, analysis must work with interconnect estimates based on available information. This is described in Section 7.

6.1 Timing optimization

This Section considers the possible savings in total path delay that could be expected by optimizing a critical path using each of the actions outlined above. In general, the savings can vary widely and are highly dependent on factors such as algorithm, design size, design characteristics, tuning parameters, etc. The numbers reported in Table 1 reflect the maximum range that one could hope to achieve with a typical design.

The third column indicates how likely a fix to one timing path comes at the expense of another timing path. A high impact indicates that savings are reduced if the critical path density is high.

We observe that placement and driver sizing and buffering offer significant savings. The logic restructuring value presented here does not include the effects of logic synthesis itself, and it is recognized that logic synthesis itself has a clear impact on timing.

7 Parasitic Models

An interconnect parasitic model is needed for any type of delay, noise or power analysis. Varying degrees of information are available during different stages of the design process. The different types of parasitic models are outlined below:

- *Fanout based parasitic model:* This model, commonly known as the wireload model in logic synthesis tools, estimates the total net capacitance and resistance based on the number of connections on a net, and on the estimated overall block size. Capacitance and resistance values are typically derived from a statistical analysis of nets in previous designs. The estimates tend to be inaccurate—for large blocks, errors can be as high

as 10x—since the model has no way of knowing whether connected cells will end up being placed adjacent to each other or on opposite corners of the block.

The model is very easy to compute and can be computed during any stage of the design process.

- *Placement based parasitic model:* This model uses placement information of connection points on a net. Particular routing directions (e.g., up-then-across versus across-then-up) may be wrong, but point-to-point distances are quite accurate. The model uses average values of capacitance and resistance *per length*. The model is more expensive to compute but more accurate than the fanout based model. Typical errors are bounded within 20%.

- *Global route parasitic model:* A global route image of all nets is required for this model. The global route data contains a coarse routing map (with known routing directions) for each net. This model understands congestion and may know, for example, the number of lines routed through a region about ten lines wide. Unlike the previous two models, this model estimates coupling capacitance more accurately, since congestion numbers are known. This is an important accuracy improvement since coupling capacitance can account for 50-70% of total wire capacitance. Typical errors with this model rarely exceed 10%.

  It is slower to compute, and can only be done after all placement and global routing is done.

- 3D extraction model. This model, described in Section 4.1 can only be calculated after full routing details are known. The model has known, reliable accuracy, but is expensive to compute.

8 Conclusions and Acknowledgements

This tutorial described the importance of, as well as methods for, extraction at the verification stage and the synthesis stage. We described both methodologies and numerical techniques for dealing with the analysis.

This work was supported by the National Science Foundation under grant CCR-97-34216, a gift from the IBM Corporation, and the DARPA Composite CAD program.

References

[1] W. C. Elmore. The transient response of damped linear networks with particular regard to wideband amplifiers. *Journal on Applied Physics*, 19(1):55–63, 1948.

[2] Jorge Rubistein, Paul Penfield Jr., and Mark Horowitz. Signal Delay in RC Tree Networks. *IEEE Trans. CAD*, 2(3):202–211, July 1983.

[3] L. T. Pillage and R. A. Rohrer. Asymptotic Waveform Evaluation for Timing Analysis. *IEEE Trans. CAD*, 9(4):352–366, April 1990.

[4] C. L. Ratzlaff and L. T. Pillage. RICE: Rapid interconnect circuit evaluation using AWE. *IEEE Trans. CAD*, 13(6):763–776, June 1994.

[5] George A. Baker Jr. *Essentials of Padé Approximants*. Academic Press, New York, NY, First edition, 1975.

[6] Curtis L. Ratzlaff, Satyamurthy Pullela, and Lawrence T. Pillage. Modelling the RC-interconnect effects in a hierarchical timing analyzer. In *Custom Integrated Circuits Conference*, pages 15.6.1 – 15.6.4, 1992.

[7] K. L. Shepard, S. Carey, E. Cho, B. Curran, R. Hatch, D. Hoffman, S. McCabe, G. Northrop, and R. Seigler. Design Methodology for the G4 S/390 Microprocessors. *IBM Journal of Research and Development*, 21(4/5):515 – 548, 1997.

[8] K. L. Shepard, V. Narayanan, P. C. Elmendorf, and Gutuan Zheng. Global Harmony: Coupled Noise Analysis for Full-Chip RC Interconnect Networks. In *Proceedings of the IEEE International Conference on Computer-Aided Design*, 1997.

[9] Peter Feldmann and Roland W. Freund. Reduced-order modeling of large linear subcircuits via a block lanczos algorithm. In 32nd *ACM/IEEE Design Automation Conference*, pages 474–479, San Francisco, California, June 1995.

[10] L. Miguel Silveira, Mattan Kamon, and Jacob K. White. Efficient reduced-order modeling of frequency-dependent coupling inductances associated with 3-d interconnect structures. In *Proceedings of the* 32nd *Design Automation Conference*, pages 376–380, San Francisco, CA, June 1995.

[11] A. Odabasioglu, M. Celik, and L. Pileggi. PRIMA: Passive reduced-order interconnect macromodeling algorithm. *IEEE Trans. CAD*, August 1998.

[12] K. L. Shepard, V. Narayanan, and R. Rose. Harmony: Static noise analysis for deep-submicron digital integrated circuits. *IEEE Trans. CAD*, pages 1132–1150, August 1999.

[13] F. Dartu and L. T. Pileggi. Calculating worst-case gate delays due to dominant capacitance coupling. In *Proceedings of the IEEE/ACM Design Automation Conference*, pages 46 – 51, June 1997.

[14] P. D. Gross, R. Arunachalam, K. Rajagopal, and L. T. Pileggi. Determination of worst-case agressor alignment for delay calculation. In *Proceedings of the IEEE/ACM International Conference on Computer Aided-Design*, pages 212 – 219, 1998.

[15] D. Edelstein and et al. Full copper wiring in a sub-0.25 μm CMOS ULSI technology. In *Proceedings of the IEDM*, pages 773–776, 1997.

[16] B. Krauter, S. Mehrotra, and V. Chandramouli. Including inductive effects in interconnect timing analysis. In *Proceedings of the CICC*, pages 445 – 452, 1999.

[17] K. L. Shepard and Z. Tian. Return-limited inductances: a practical approach to on-chip inductance extraction. In *Proceedings of the CICC*, pages 453 – 456, 1999.

[18] M. W. Beattie and L. T. Pileggi. IC analyses including extracted inductance models. In 36th *ACM/IEEE Design Automation Conference*, pages 915 – 920, 1999.

[19] P. Restle, A. Ruehli, and S. G. Walker. Dealing with inductance in high-speed chip design. In 36th *ACM/IEEE Design Automation Conference*, pages 904 – 909, 1999.

[20] H. H. Chen. Minimizing chip-level simultaneous switching noise for high-performance microprocessor design. In *Proceedings of the IEEE International Symposium on Circuits and Systems*, volume 4, pages 544 – 547, 1996.

[21] Umakanta Choudhury and Alberto Sangiovanni-Vincentelli. Automatic generation of analytical models for interconnect capacitances. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 14:470 – 480, April 1995.

[22] F. Preparata and J. Nievergelt. Plane-sweep algorithms for intersecting geometric figures. *Communications of the ACM*, 25(10):739 – 747, 1982.

[23] T. G. Szymanski and C. J. Van Wyk. Goalie: a space-efficient system for VLSI artwork analysis. *IEEE Design and Test of Computers*, 2(3):64 – 72, June 1985.

[24] K.-W. Chiang, S. Nahar, and C.-Y. Lo. Time-efficient VLSI artwork analysis algorithms in GOALIE2. 25th *ACM/IEEE Design Automation Conference*, pages 471 – 475, 1988.

[25] U. Lauther. An O(nlogn) algorithm for Boolean mask operations. 18th *ACM/IEEE Design Automation Conference*, pages 555 – 562, 1982.

[26] K. P. Belkhale and P. Banerjee. Parallel algorithms for VLSI circuit extraction. *IEEE Trans. CAD*, 10(5):604 – 618, 1991.

[27] A. Farrar and A. T. Adams. Matrix methods for microstrip three-dimensional problems. *IEEE Transactions on Microwave Theory and Techniques*, MTT-20(8):497–504, August 1972.

[28] A. E. Ruehli. Equivalent circuit models for three-dimensional multiconductor systems. *IEEE Transactions on Microwave Theory and Techniques*, MTT-22:216 – 221, March 1974.

[29] M. Kamon, N. Marques, Y. Massoud, L. Silveira, and J. White. Interconnect analysis: From 3-d structures to circuit models. In *Proceedings 36th Design Automation Conference*, pages 910–914, New Orleans, La, June 1999.

[30] M. Kamon, M. J. Tsuk, and J. White. Fasthenry: A multipole-accelerated 3-d inductance extraction program. *IEEE Transactions on Microwave Theory and Techniques*, 42(9):1750–1758, September 1994.

[31] Nuno Marques, Mattan Kamon, Jacob White, and L. Miguel Silveira. An efficient algorithm for fast parasitic extraction and passive order reduction of 3d interconnect models. In *Proceedings of DATE'98, Design Automation and Test in Europe*, February 1998.

[32] M. Kamon, N. Marques, L. M. Silveira, and J. White. Automatic generation of accurate circuit models of 3-d interconnect. *IEEE Transactions on Components, Packaging, and Manufacturing Technology-Part B*, 21(3):225–240, August 1998.

[33] Jing-Rebecca Li, Frank Wang, and Jacob White. An efficient lyapunov equation-based approach for generating reduced-order models of interconnect. In *Proceedings 36th Design Automation Conference*, pages 1–6, New Orleans, La, June 1999.

[34] B.-T. Lee, E. Tuncer, and D. P. Neikirk. 3-d series impedance determination using a surface ribbon method. In *Proceedings of the* 4th *Topical Meeting on Electrical Performance of Electronic Packaging*, pages 220–222, Portland, Oregon, October 1995.

[35] J. Wang, J. Tausch, and J. White. A wide frequency range surface integral formulation for 3-d inductance and resistance extraction. In *International Conference on Modeling and Simulation of Microsystems, Semiconductors, Sensors and Actuators*, San Juan, Puerto Rico, April 1999.

[36] Eli Chiprout, Hansruedi Heeb, Michael S. Nakhla, and Albert E. Ruehli. Simulating 3-D retarded interconnect models using complex frequency hopping (CFH). In *Proceedings of the IEEE/ACM International Conference on Computer Aided-Design*, pages 66–72, Santa Clara, California, November 1993.

[37] J. R. Phillips, E. Chiprout, and D. D. Ling. Efficient full-wave electromagnetic analysis via model-order reduction of fast integral transforms. In *Proceedings 33rd Design Automation Conference*, Las Vegas, Nevada, June 1996.

[38] W. E. Donath et al. Timing driven placement using complete path delays. In *Proceedings of the 27-th ACM/IEEE Design Automation Conference*, pages 84–89, June 1990.

[39] A. B. Khang, S. Muddu, E. Sarto, and R. Sharma. Interconnect tuning stratagies for high-performance ics. In *IEEE/ACM Design, Automation and Test Conf. in Europe*, 1998.

[40] E. M. Sentovich et al. Sis: A system for sequential circuit synthesis. Technical report, Electronics Research Laboritory, University of California, Berkeley, May 1992. Memo No. UCB/ERL M92/41.