Optimal Interconnects: Modelling and Synthesis

Vineet Sahula[‡] C.P. Ravikumar D. Nagchoudhuri
Department of Electrical Engineering
Indian Institute of Technology. New Delhi, INDIA.
{sahula,rkumar,dnag}@ee.iitd.ernet.in

ABSTRACT

The arrival of deep submicron technologies (DSM) has opened up a new dimension to all problems related to VLSI design and test. These technologies allow us to build entire systems on a chip, and as a result, the complexity of problems such as logic synthesis, technology mapping, and physical design has gone up tremendously. More importantly, some of the assumptions related to the effect of interconnect delays are no more valid in the submicron domain. Interconnect is now responsible for 90% of signal delays in a DSM integrated circuit. As a result, problems such as circuit partitioning, placement, and routing must be predominantly concerned with the minimization of interconnect delays. Problems at the higher level of design abstraction, such as high-level synthesis and logic synthesis must also address the issue of interconnect delays. VLSI engineers now agree that failure to take interconnect-related issues into account early in the design life cycle can lead to a failure in meeting the specification, leading to costly design iterations. In this paper, our objective is to conduct a survey of state-of-the-art papers in modeling and synthesis of interconnects.

Indexing Terms : *Deep submicron, Distributed RLC networks, Time moments, Padé approximation, transmission line synthesis, driver sizing, clock-tree routing.*

1. INTRODUCTION

The design of integrated circuits is again being transformed due to the increase in circuit density, increase in chip dimensions, smaller feature sizes and higher operating frequencies.

Simple lumped RC approximations are often insufficient and RLC transmission-line model becomes necessary in the analysis of modern-day ICs We begin by identifying the need for more detailed models. We then introduce the models for interconnect delay which have been proposed in the literature (Section 2). We also discuss the techniques to transform the transfer function of a transmission line into a reduced order rational function through the *reciprocal expansion approach* using Padé approximation. Section 3 presents a comparative discussion of various approaches for interconnect synthesis. The objectives of interconnect synthesis are (i) Delay optimisation, (ii) Area optimisation, and (iii) Power optimisation.

1.1 Interconnects dominate the modern design process

The word *Giga Scale Integration* (GSI) is being used to describe levels of integration beyond 10^9 /chip [22]. Interconnection is the most crucial contributing factor to IC performance, since wire delays are responsible for 90% of signal delays in a DSM chip [3]. Routing of power lines and clock lines has always been an area of special emphasis in VLSI physical design. With DSM geometries, the nonlinearity and the distributed nature of the interconnect complicate the picture.

In the traditional design approach, a designer spends more time in the higher levels of design abstraction, progressing from behavioral level through register-level-transfer (RTL) to logic level. Physical design is the last phase in the design process, therefore, the effect of interconnections is not modelled by the front-end tools.. Design engineers agree that physical design should receive primary consideration when DSM geometries are involved, to ensure that interconnection information is available throughout the design process. Alternately, one may *predict* wiring delays accurately at an early stage in logic design.. Thus both analysis and synthesis of interconnect have emerged as problems of paramount importance in the recent past.

To illustrate the interplay between physical and logic design, consider the following scenario. Designers often increase the speed of a system by identifying the critical path and reducing the delay of logic blocks on the critical path through techniques such as logic decomposition. Buffers are then inserted along the critical path so as to employ pipelined operation. Estimation of critical delay has traditionally ignored the wiring delays. Architectural optimisations such as pipelining are difficult to perform in the DSM domain, since estimating critical delays at an early stage of design is not easy.

[‡] V.Sahula is with the Department of Electronics & Comm. Engg., Malaviya Regional Engineering College, JAIPUR. He is currently a Ph.D candidate at IIT,DELHI.

Table 1 : Modelling of Interconnects

Condition	Model Applicable
$t_f > 0.4* t_r$	Transmission line phenomenon becomes significant
$t_f < 0.2* t_r$	Lumped capacitor model is sufficient.
$0.4*t_r > t_f > 0.2* t_r$	Transmission-line model, distributed-RC model or lumped capacitance approximation depending on required accuracy

Solutions to the above problem can be (i) prediction of *likely* interconnect routes as a guide to circuit layout (ii)an incremental approach to design layout and verification so that errors may be corrected at the same time when more detailed information is being developed.

.Fast place & route tools and delay prediction tools play an important role in the above context. Physical design tools, that are capable of integrating floorplanning, placement, routing, and signal analysis can deliver estimates of interconnect effects useful to the front-end design tools. Such preliminary interconnection data can then be passed to the back-end design tools to serve as constraints on later physical design processes. Thus we need

(i) fast and *incremental* place & route,

(ii) incremental logic optimisation to address physical problems, and

(iii)detailed optimisation of the interconnect itself.

2. INTERCONNECT MODELLING AND DELAY ESTIMATION

2.1 Transmission Line model :

Traditionally interconnects have been treated as lumped capacitive loads. As feature size becomes smaller interconnect resistance gains significance and the contribution of the fringing capacitance to the total interconnect capacitance becomes larger. Interconnects must then be analysed as distributed RC-networks. With chip size becoming larger and system speeds becoming higher, interconnect lengths have begun to approach signal wavelength; in such situations, the inductance of the interconnect also becomes important and the wires must be modelled as *RLC transmission lines*. This is the case with chip-to-chip interconnections on multichip modules as well as printed wire boards.

Transmission line behavior becomes significant when the time of flight t_f becomes comparable to signal's rise time t_r , defined as the time taken by a signal to rise from 10% to 90% of its final value See Table 1. The time of flight may be expressed as

$$t_f = \frac{l}{v}$$

where l is the line length and v is the velocity of wave propagation. If wire resistance is greater than the characteristic impedance of the line, then resistive effect dominates the inductive phenomenon, and a wire can be modelled as a distributed RC-line instead of RLC-transmission line.

2.1 Rapid Interconnect Simulation

Several techniques have been developed for the simulation of the transient response of a lossy transmission line. However, a typical simulation for a single net using these techniques can be very time-consuming[20].

Sectioned, Lumped Models : To avoid the direct solution of equations related to transmission lines, we can logically partition a wire into several *sections* and model each section of the wire as a lumped RLC circuit. By selecting a sufficiently large number of sections, we can ensure that the distributed properties of interconnect lines are preserved. However, circuit simulation or timing analysis are still very time consuming due to the large number of sections.

We now discuss some of the model approximation techniques found in the literature. The transfer function for an RLC tree with N sections of unit length is a $2N^{th}$ order function

$$H(s) = \frac{a_0 + a_1 s + a_2 s^2 + \cdots + a_{2N-1} s^{2N-1}}{b_0 + b_1 s + b_2 s^2 + \cdots + b_{2N} s^{2N}}$$

An approximation to H(s) can be found by computing the *moments* of either H(s) or G(s) = 1/H(s). The moments of a transfer function are defined to be the coefficients of the Maclaurin series expansion of the transfer function.

$$H(s) = m_0 + m_1 s + m_2 s^2 + \ldots + m_k s^k + \ldots$$

$$\text{where} \quad m_k = \frac{(-1)^k}{k!} \int_0^\infty t^k h(t) dt$$

The moment series is converted into a rational function **H(s)** of lower order using Padé approximations [4]. This technique is widely known as *Asymptotic Waveform Evaluation*. An equally popular technique, known as *Reciprocal Expansion* (or REX) [11] which approximates G(s). The poles of **H(s)** generated through Padé approximation correspond to the dominant poles of the original system as well as a few poles that do not correspond to poles in the original system, but account for effect of the remaining poles. Table 2 compares several approximation techniques [4,5,6,7,21].

"Krylov" methods reported in [5] are based on the implementation of Padé approximation through the use of *Lanczos process* (Padé Via Lanczos). Some techniques use a direct simulation of transmission line[20].

Table 3. Second order delay estimation

Conditions	Delay Estimate	Description of time-domain output for step input
$K < 1$	$t_d = b_1$	Monotone rising
$1 \le K \le (4\pi^2 + 1)$	$t_d = [1 + \frac{1}{2}\sqrt{K-1}]\, b_1$	Oscillation with undershoot not falling below 63% of final value
$K \ge (4\pi^2 + 1)$	$t_d = K b_1 / 2$	Oscillation with undershoot falling below 63% of final value

Table 4(a) : Delay Estimates for Ramp Input [9]

$$V_{in}(s) = \frac{V_o}{T_R}.\frac{1}{s^2}[1 - e^{s T_R}]$$

$$V_{out}(s) = \frac{V_o}{T_R}.\frac{1}{s^2}[1 - e^{s T_R}].H(s)$$

Hence $V'_{out}(s) = \int_0^\infty V'_{out}(t) - \int_0^\infty t\, V'_{out}(t)dt$

$$= V_o\left(1 - \frac{s T_R}{2} + \ldots\right)\left(\frac{1 + a_1 s + a_2 s^2 + \ldots}{1 + b_1 s + b_2 s^2 + \ldots}\right)$$

Therefore $T_{AR} = \frac{T_R}{2} + b_1 - a_1 = \frac{T_R}{2} + T_{ED}$

where
V_{in} : input ramp voltage with magnitude V_o and rise time T_R
V_{out}, V'_{out} : output voltage and its first derivative

T_{ED} : Elmore delay under unit step input
T_{AR} : delay estimate under ramp input

Table 4 (b). Voltage response of an open-ended distributed RC line under a step input excitation [23]

Method	*Accuracy.*	*Time-domain voltage response*
Simple lumped	approximate	$V_0(1 - e^{-\frac{t}{RC}})$
Distributed RC (Wilnai's two-port)	small t	$2\, V_0(1 - erf\sqrt{\frac{RC}{4t}})$
	large t	$V_0(1 - 1.366\, e^{-2.5359\frac{t}{RC}} + 0.366\, e^{-9.4641\frac{t}{RC}})$
Diffusion equation	exact	$V_0(1 - erf\sqrt{\frac{RC}{4t}})$

2.2 Delay Models

The *Elmore* time-constant model, a first order model commonly used for lossy interconnect, takes into account the resistance of the interconnection and is more accurate than a lumped capacitance model. For an RC tree which approximates the interconnection, the delay from the input to a node e can be computed as :

$$T_e = \sum_k R_{ke} C_k$$

R_{ke} is the resistance of the path common between a node *k* to the input and the node *e* to the input. C_k is the capacitance at node *k*. Bounds for step response delay based on a *one pole approximation with time-constant* can be determined using circuit element values [8]. However, when transmission line effects become significant, this model is of limited use since it ignores the inductance of interconnection, and cannot predict oscillations in the waveform at the receiver end of the line. A reasonable compromise between accuracy and speed is to use a *second-order delay model*. The following second order function represents the transfer function from a driver node to a receiver node :

$$H(s) = \frac{1}{b_0 + b_1 s + b_2 s^2}$$

By defining $K = 4b_2 / b_1^2$, the delay bound can be computed as shown in Table 3. Table 4(a) provides the delay estimates for *ramp* input. Table 4(b) provides the time domain response for unit step input.

3. Synthesis

In deep submicron design, the problem areas for the interconnect-layout optimisation-algorithms are optimal topology selection, optimal wire sizing, routing (clock tree generation), gate(driver) sizing and fanout optimisation etc. Various algorithms have been reported for these problems[12-20].For performance-optimisation, constraining variables are time, area and power dissipation. EDA programs being developed are based on the concepts of performance-driven-routing and Layout-driven synthesis.

In deep submicron, accuracy of propagation delay interconnect is critical to the design of high speed and high performance systems. Existing wire-delay estimation techniques are based either on simulation or (closed form) analytical formulae. Simulation methods (SPICE) give very accurate insight into arbitrary interconnect structure but are computationally very CPU expensive. Faster methods using moment matching techniques are available [4][5][7](as have been discussed in section 2) but they are also expensive to be used during interconnect synthesis & optimisation. A first order approximation of Elmore delay under step input has been most widely used analytical estimate for interconnect delays during interconnect synthesis and routing. The algorithms developed for VLSI routing are inadequate for interconnect routing, they try to minimise the wire length for minimising delay (this is true with the lumped-capacitor model for wire). Whereas for

interconnect routing, when Elmore or second order delay model is used, minimising wire length does not necessarily imply minimum interconnection delay. We have compared various routing algorithms in Table 5.

4. Conclusions & Future trends

Thus we have surveyed the interconnection-design issues, problems and their probable and available solutions. As a future prediction, the EDA paradigm will be required to be changed. High level or Logic level designs based on conventional estimated delay are no longer useful and layout-driven-synthesis & optimisation concept is required. With sub-half-micron technologies creating a fundamental shift in the problem faced by IC designers, coupling effects signal integrity & noise are assuming a critical role in determining IC performance. The correct modelling of coupling effects is also a major challenge in deep submicron VLSI design. The coupling effects can be significantly minimised but not completely eliminated through proper design methodology and it is impossible to adequately simulate coupling effects in large VLSI designs. There is also a necessity to develop efficient physical design algorithms and tools to control the noise maintaining signal integrity during layout design process, with the same emphasis as on area, delay and power optimisation. This requires first developing accurate yet efficient noise models to be incorporated into the layout tools. Based on these models, existing layout optimisation techniques for delay minimisation, such as routing topology optimisation, buffer insertion, driver & interconnect sizing shall be extended to address the noise and signal integrity issues.

References

[1] H.B. Bakoglu. *Circuits, Interconnections and Packaging for VLSI*, Addison-Wesley,1990.

[2] A.K.Goel. *High Speed VLSI Interconnects*, Wiley, 1994.

[3] T.Chang Lee and J.Cong, "The new line in IC design'', *IEEE Spectrum*, March 1997, pp. 52.

[4] L.Pillage and R.Rohrer, "Asymptotic Waveform Evaluation for Timing Analysis'', *IEEE Trans. CAD*, vol. CAD-9, April 1990, pp 352-366.

[5] P.Feldmann and R.W.Freund, "Efficient linear circuit analysis by Pade Via Lanczos process'', *IEEE Trans. CAD*, vol. 14, May 1995, pp. 639-649.

[6] A.J. van Genderen and N.P. van der Meijs, "Reduced RC models for IC Interconnections with coupling capacitances,'' *Proc. IEEE 3rd EDAC.*, March 1992, pp. 132-136.

[7] P.J.S.Elias and N.P.van der Meijs, "Extracting Circuit Models for Large RC Interconnections that are Accurate upto a Predefined signal Frequency,'' *Proc. 33rd DAC.*, 1996

[8] B.Tutuianu, F.Dartu and L.Pileggi, "An Explicit RC-Circuit delay Approximation Based on First Three Moments of Impulse Response,'' *Proc. 28th DAC.*, June 1991, pp. 555-560.

[9] A.B.Kang, K.Masuko and S.Muddu, "Analytical Delay Models for VLSI Interconnects Under Ramp Input,'' *Proc. ICCAD*, 1996

[10]C.L.Ratzlaff, N.Gopal and L.T.Pileggi, "RICE : Rapid Interconnect Circuit Evaluator,'' *Proc. 28th DAC.*, June 1991, pp. 555-560.

[11]M.Sriram and S.M.Kang, "Fast Approx. of the Transient Response of Lossy Transmission Lines Trees'', *Proc. 30th DAC.*, June 1993, pp. 691-696.

[12]J.Cong, S.Leung and D.Zhou, "Performance-Driven Interconnection Design Based on Distributed RC Delay Model,'' UCLA CS Dept. Tech. Report CSD-920043, Los Angeles, Oct. 1992.

[13]D.Lehther and S.Sapatnekar, ``Clock Tree Synthesis for MCMs,'' *Proc. 33rd ICCAD*, 1996

[14]R.S. Tsay, `` An Exact Zero Skew Clock Routing Alg.,'' *IEEE Trans. CAD*, v-12, 1993, pp. 242-249.

[15]R.Gupta, B.Krauter and L.Pileggi, "Transmission line Synthesis via Constr. Multivariable Opt.'' *IEEE Trans. CAD*, vol. CAD-1, Jan. 1997, pp 6.

[16]K.Sato et al. "Post-Layout Optimisation for Deep Submicron Design,'' *Proc. 33rd Design Automation Conf.*, June 1996.

[17]J.Cong and C.K.Koh, " Simult. Driver and Wire Sizing for Performance and Power Optimisation,'' *IEEE Trans. VLSI*, 2(4), Dec. 1994, pp. 408-425.

[18]T.Okamoto and J.Cong, " Buffered Steiner Tree Construction with Wire Sizing for Interconnect Layout Optimisation,'' *Proc ICCAD*, 1996.

[19]C.P.Chen, H.Zhou and D.F.Wong, "Optimal Non-Uniform Wire-Sizing under the Elmore Delay Model,'' *Proc. ICCAD*, 1996.

[20] E.C.Chang, S.M.Kang, "Transient Simulation of lossy coupled transmission lines using Iterative Linear Least Square fitting and Piecewise Recursive Convolution,'' *IEEE Trans. CAS-I*, 43(11), 1996

[21] K.Nabors et al. "Lumped interconnect model via Gaussian Quadrature,"*Proc. 34th DAC.*,June 1997.

[22] "Giga Scale Integration", *IEEE Spectrum 1996*

[23] A.B.Kang and S. Muddu, "Delay analysis of VLSI interconnections using the Diffusion Equation model,"*Proc 31st DAC.*, June 1994.

Table 2. A comparison of Interconnect modelling techniques

Method	*Accuracy/Regime*	*Methodology adopted*	*Computational method*	*Advantages*	*Disadvantages*
AWE[4]	Time moments matching	1.Moments series generated from transfer-function 2.A reduced order transfer-function obtained from moment-series(Padé Approximation) 3.Macro models obtained from poles	Matrix Algebra	Reduces number of system poles greatly	1.Not always numerically stable 2.Circuit model must be constructed from the poles obtained
EPR[6]	Fairly accurate at low frequencies	1.Preserves Elmore delay during reduction 2.Yields complete input/output relation between terminal nodes		1. Simple model 2. Numerically stable for RLC lines	1. Removes all non-terminal nodes; original circuit structure lost 2. Not valid for high frequencies as it provides only 1st order approximation
Modified EPR [7]	Accurate upto user defined frequency f_s	1. Selectively removes non-terminal nodes removal of which cause tolerable error 2. Provides reduced terminal equivalent of original network	1.Estimates relative error made if a node is eliminated 2. Relative error computed directly in terms of moments of admittance function & frequency f_s	1. High accuracy upto f_s 2.Network topology nearly resembles original network. Hence physically realisable 3.Numerically stable	1. Applicable to RC circuits, method not well established for RLC nets 2. If extended to RLC networks accuracy is bound to suffer as it is based on Elmore delay preservation
Gaussian Quadrature [21]	Moment-matching of circuit-element distribution-function, not time moment-matching	1.L and C of line described as functions of line-resistance R 2.L and C expressions converted into weighted sum and required moments matched 3.Preserves Elmore delay for RC-networks	Gaussian Quadrature method used for appropriating integrals by weighted sums	1. Stable models 2. Versatile(can be used to compact RC, RLC or LC) 3. Conversion to time moments possible	.Method works better with RC & LC lines only
PVL[5]	1. Same computa-tional complexity as AWE 2. Time moment-matching	Exploits good relation between Lanczos process & Padé Approximation	Iterative	Capable of estimating accuracy of reduced order model obtained by it	1.Large no. Of iterations may be required 2.Stability for RLC networks not always assured

Table 5. Interconnect optimisation and synthesis techniques : A comparison

Method	*Methodology/ description*	*Topology selection*	*BI*[#]	*WS*[#]	*DS*[#]	*LT*[#]	*Delay model used*	*Specific remarks*
Interconnect synthesis for MCMs [13]	1.Zero skew clock tree construction 2.REX for transfer function approximation 3.Lines are critically damped by series termination	✓	✓	×	×	✓	Distributed RLC, 2nd order model	
Post-layout optimisation [16]	1.Wire RC information is back annotated to logic synthesis system 2.Buffer-insertion & gate-sizing are applied as post-layout optimisation 3.Rerouting using engineering-change-order to preserve previous layout structure	×	✓	×	✓	×	1st order Elmore delay model	Provides simultaneous buffer-insertion & driver-sizing
Interconnect layout optimisation [12]	1.Fanout optimisation using concepts of critical path isolation & balanced load decomposition 2.Steiner tree construction 3.Dynamic programming	✓	✓	✓	×	×	Elmore delay model of I-order for interconnects, RC delay model for buffer	Provides simultaneous buffer-insertion & wire-sizing
Delay & power optimisation [17]	1. Formulated as constrained optimisation problem given a routing tree	×	×	✓	✓	×	Distributed Elmore delay RC delay model for driver	Provides simultaneous driver and wire sizing with two objectives : - delay minimisation only - combined delay & power minimisation
Non-uniform wire-sizing [19]	1. Optimisation objectives : -maximize. delay -minimize area subject to delay bound 2.Greedy algorithm 3. Constrained non-linear optimisation solved by Lagrangian relaxation method	×	×	✓	×	×	Distributed Elmore delay model	Non-uniform wire sizing i.e. single wire can have varying width along its length given by distribution function ae^{-bx} for $0 \leq x \leq L$. Wire width at position x is ae^{-bx}
RLC interconnect synthesis [15]	1. Formulated as constrained multivariables optimisation 2.Takes into account various factors i.e. loading & loss in line and rise time of input signal 3.Moments based line-termination 4.Use of moments-sensitivities (in determination of moments)	✓	×	✓	×	✓	Distributed RLC delay model	1. Different design variables can be considered concurrently e.g. width, resistivity of interconnect, resistive source or far-end termination 2.Trade between signal rise-time & ringing time possible without requiring time domain solution

[#] *BI* : Buffer Insertion *WS* : Wire Sizing *DS* : Driver Sizing *LT* : Line Termination