

Measurement and Modeling of On-Chip Transmission Line Effects in a 400 MHz Microprocessor

P. J. Restle, K. A. Jenkins, A. Deutsch, and P. W. Cook

Abstract—On-chip interconnect delays are becoming an increasingly important factor for high-performance microprocessors. Consequently, critical on-chip wiring must be carefully optimized to reduce and control interconnect delays, and accurate interconnect modeling has become more important. This paper shows the importance of including transmission line effects in interconnect modeling of the on-chip clock distribution of a 400 MHz CMOS microprocessor. Measurements of clock waveforms on the microprocessor showing 30 ps skew were made using an electron beam prober. Waveforms from a test chip are also shown to demonstrate the importance of transmission line effects.

Index Terms—Clock distribution, clock distribution networks, clock trees, e-beam probing, electron-beam probing, high-speed measurement, inductance, microprocessor chip, transmission lines, VLSI circuits.

I. INTRODUCTION

DUE TO increasing chip size, more function on a chip, and higher clock frequencies, on-chip wiring is having a greater impact on microprocessor performance. Since on-chip wiring delays are a significant fraction of some cycle-time limiting paths, it is becoming more important to reduce wiring delays, and to model wiring delay more accurately. A special case is the on-chip clock distribution, where one or more clock signals must be distributed simultaneously to virtually all regions of the chip, with accurately known delays. Unexpected differences in the clock distribution delays (skew) will reduce the allotted cycle time and, if large enough, can cause functional errors.

For CMOS microprocessors, short on-chip wires can often be modeled as simply an additional parasitic capacitive load (C wire model). For most medium length on-chip interconnects a model including distributed wire resistance and capacitance (RC model) results in acceptable accuracy. However, it will be shown that for certain critical cases with long wires needing greater than minimum wire widths, such as aggressively designed on-chip clock distributions, transmission line effects (LRC wire model) become very significant.

Although the growing importance of on-chip transmission line effects has been predicted [1], the lack of published evidence from microprocessor designs, and the difficulty of modeling and simulating lossy, nonuniform transmission lines using existing CAD design, simulation, and timing tools has prevented consideration of these effects in most cases. One exception is the clock distribution described in [2], where

microstrip transmission lines were used for the first-level wiring tree of the clock distribution. That technique provides simple extraction and control of transmission line parameters, but requires significant disruption of the wiring on two metal layers to implement the ground return. Another strategy is to add dedicated metal layers to the fabrication process as V_{DD} and ground reference planes. This also greatly simplifies analysis of transmission line parameters [3], but adds to the chip cost (unless these planes are also required for on-chip power distribution).

This paper describes a clock distribution with measured 30 ps skew on a 400 MHz S/390 CMOS microprocessor [4]. This was achieved using less than 2% of the top two metal layers (M4 and M5), with only the M5 level being a low resistance level. The technology and circuit details are described more fully in [4]. This low-skew clock distribution did not disrupt the usual wiring methodology: all horizontal clock wires and shields were restricted to M5, and vertical wires were on M4. Neither microstrip structures [2] nor on-chip reference planes [3] were used to simplify the transmission line analysis. The simulations and measurements show significant but controlled transmission-line effects. This is the first report of measurements on a microprocessor chip having sufficient accuracy to confirm the validity of on-chip transmission line models. These results are contrasted below with results from a test chip which was designed without consideration of transmission-line effects.

II. CLOCK DISTRIBUTION AND MODELING

A single clock is globally distributed from a centrally located on-chip phase-lock-loop (PLL) through a central chip buffer, to 580 distribution points (clock pins) on functional blocks (macros). This global clock distribution from the central buffer to 580 clock pins is the focus of this paper. The distribution is achieved in two levels of balanced H-like trees. The first and second levels of the clock distribution are shown in Fig. 1. The first-level tree routes the global clock from the central clock buffer to the nine sector buffers, with each buffer consisting of three inverters in series. The nine sector buffers repower the clock to a total of 580 macro clock pins within the processor units. These first two levels of clock distribution used primarily the top two metal layers. Due to the asymmetric distribution of clock loads and widely varying capacitances of individual clock pins, a clock wiring methodology was developed with special purpose routing, timing, and tuning CAD tools developed by some of the authors. With these tools, the detailed routing, as well as the widths of all clock wires, were optimized to minimize skew, mean delay, power,

Manuscript received April 1, 1997; revised September 5, 1997.

The authors are with the IBM T. J. Watson Research Center, Yorktown Heights, NY 10598 USA.

Publisher Item Identifier S 0018-9200(98)02340-3.

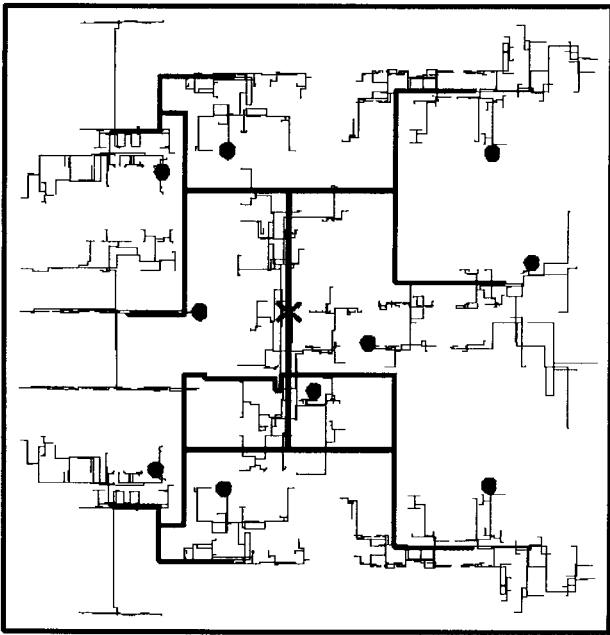


Fig. 1. CMOS S/390 microprocessor first- and second-level wiring tree, with dots showing the measurement points for waveforms in Fig. 3. The wire widths are not to scale. The wide wires are the first-level wiring tree driving nine sector buffers, and the narrow lines are the second-level trees.

wiring tracks, and sensitivity to process variations. The choice of a small number of large buffers each surrounded by on-chip decoupling capacitors also reduces skew and jitter from on-chip process and power variations, but results in more complicated wiring networks. In order to obtain the fastest possible distribution, the techniques of adding capacitance, or padding, and increasing wire lengths with serpentes were not used. Distributed LRC modeling, described below, was used for virtually every wire in the trees during the design (routing and tuning) processes.

A transmission line model must take into account the wire inductance (L), wire capacitance (C), as well as resistive (R), losses. Chip-to-chip interconnects are relatively uniform lower-loss transmission lines. On-chip wiring can often be adequately represented by lumped C or RC models. However, when the transition time of a signal becomes comparable to its propagation delay, transmission line effects must be represented with distributed LRC parameters [1]. This will occur when wide metal is used to reduce resistance on long lines. On-chip interconnections have unique characteristics, namely high resistive losses and very nonuniform transmission line structures. The effective resistance model must include the resistance of the return path of the transmission line (return-path resistance), and thus depends not only on the signal wire, but also on the geometry of nearby power and signal wires [5].

Since a return path was not provided using parallel wires on adjacent layers, three-dimensional (3-D) modeling was performed using a full-wave electromagnetic field solver [6]. A large number of wiring geometries cases were analyzed, and the results were used to generate a combination of analytic models and look-up tables containing distributed LRC parameters for all clock wiring geometries used. Each wire

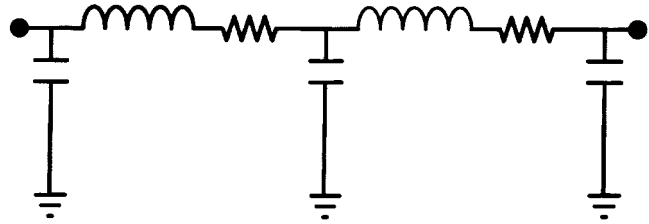


Fig. 2. Transmission line "Pi Model" with two segments.

segment of the clock distribution network was represented by an equivalent circuit consisting of up to 6 LRC "Pi" segments, as shown in Fig. 2. Finer segmentation was not needed. The effective resistance, R , of an on-chip transmission line can increase significantly with frequency, while the inductance decreases. This is not primarily due to skin-effect or the silicon substrate, but to the resistance of the nearby power bus wiring that carries the inductive return current. At higher frequencies, the return current concentrates in the nearest power (and signal) wiring. During the design process the total effective line resistance R in Fig. 2 was modeled as the DC line resistance in series with a return path resistance. A characteristic frequency was chosen so that the inductance and resistance used in the design process was independent of frequency. RICE [7] and ASTAP [8] were used for circuit simulation and wire width tuning. Including inductance increases simulation time by approximately a factor of two.

Reducing skew for the nominal parameters was only one of the design goals, and zero simulated skew was not achieved. Simulations using the best estimates for process and model parameters (i.e., nominal conditions) predict a maximum skew of 40 ps, with less than 20 ps skew for the majority of the 580 clock pins. Additional skew resulting from switching noise, process variations, and modeling inaccuracies was modeled and minimized during the design process, but was still expected to be significant. The success of this design method is shown by the measurements presented below.

III. CLOCK SKEW MEASUREMENT

A novel electron-beam prober, having a 15 ps time resolution [9], was used to measure the waveforms of the clock signal at the points on the top wiring layer indicated in Fig. 1. No special pads or test structures were required to enable electron-beam probing, since the clock distribution used the top metal layer for most horizontal wiring, and the 0.2 μm spatial resolution of the prober allowed probing of minimum width wires.

The only preparation required was the removal of solderballs and the passivation above the top wiring layer. The capacitance change due to removal of the passivation was small, but was included for the model-hardware comparisons. Due to uncontrollable local field effects on electron-beam waveform amplitudes, each measured waveform was automatically corrected by calculating a voltage offset and voltage scaling factor to match the assumed switching between Ground and V_{DD} . This assumption was checked with mechanical probing. This adjustment of the starting and ending voltage of each measured waveform were the only "fitting parameters" used.

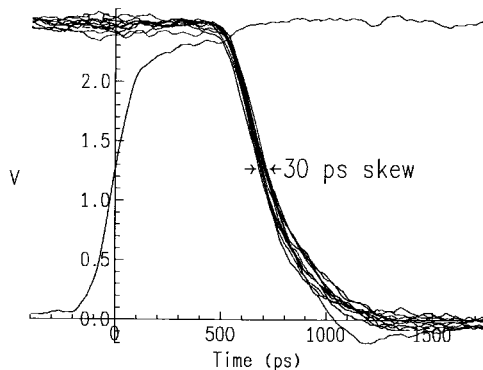


Fig. 3. Clock waveforms measured on the microprocessor at the points shown in Fig. 1, showing 30 ps skew.

Because the chip was powered using a standard cantilever probe card in the electron-beam prober, the chip clock was run at low frequency to minimize power supply noise. Power supply noise during these measurements was measured, on chip, to be less than 100 mV.

The clock signals were measured at some of the internal points described above. Fig. 3 shows ten waveforms measured at one or two randomly chosen clock pins within each of the nine sectors, indicated in Fig. 1. Because electron-beam probing requires waveform averaging, there was no measure of clock jitter. The total measured skew was 30 ps across the chip. It is noted that this skew includes both wiring and sector buffer delays. Both the first- and second-level trees showed significant transmission line effects. Although there have been designs with simulated clock skews as low as 90 ps [10] and local tuning claiming an accuracy of 10 ps [2], this is the first report of measured skew in this time scale. As will be shown below, this extremely low skew depended critically on the accuracy of the transmission line models.

IV. TEST-SITE MEASUREMENTS

A test chip with a similar clock distribution was fabricated prior to the microprocessor design. However, the test chip was designed and tuned without consideration of transmission-line effects. The significant differences between measurements on this test chip and the RC model waveforms thus illustrate the errors of ignoring on-chip transmission-line effects. Fig. 4 shows the waveforms at points on the first-level tree indicated in the insert, along with the waveforms simulated with the RC model. A modeling error of almost 200 ps, half of the measured wiring delay, is evident.

Fig. 5 shows delay measurements on another branch of the top level clock tree of the test chip. Modeled waveforms are shown for three different levels of modeling accuracy. The RC model shows the results of ignoring all transmission line effects resulting in a delay error of 65 ps. An improvement results from including the line inductance, but not considering the effect of return-path resistance. Finally, the full model obtained from the 3-D field solver, including return-path resistance, results in waveforms much closer to the measurement.

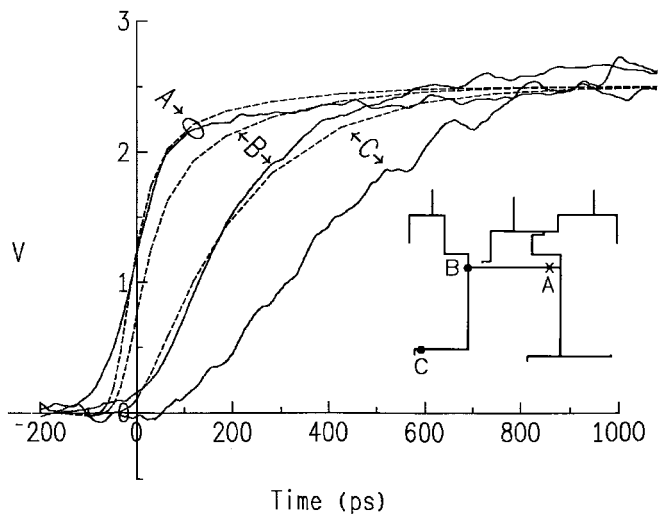


Fig. 4. Test chip waveforms (solid) with RC wire model (dashed) showing large delay error due to transmission line effects. Inset shows the first-level clock tree and the measurement points.

The largest delay error in Fig. 4 was due to omission of the inductance for the longest, widest wire on the low-resistance (thick) top metal layer (M5). Fig. 5 shows that transmission line effects can be significant even for wires on the thinner, more resistive M4 metal layer, for which an RC model might be thought to be accurate. In Fig. 5(a), it is seen that including the wire inductance improves the estimate of wiring delay significantly, but actually reduces the accuracy of the transition time. Including both the inductance and the increase in effective resistance results in good results for both the delay and transition time. Fig. 5(b) shows good agreement of the measurements with the LRC model. It is noted that the electron-beam measurements were essential for demonstrating the importance of transmission line effects, and for establishing the accuracy of the model needed to achieve low clock skew.

The total measured skew at six places on the second level trees of the test chip was approximately 150 ps, resulting primarily from transmission line effects in the design. This should be contrasted with the 30 ps skew obtained on the 400 MHz microprocessor. This much smaller skew (a factor of five) was primarily due to consideration of transmission line effects at every stage of the microprocessor design. For example, near critical lines the local power supply bus was modified to provide better transmission line characteristics. The desired characteristics included matched impedances, reduced inductance and capacitance to improve signal velocity, and lower return path resistance. In addition, geometries were chosen to improve reliability, predictability, and reduce process sensitivity.

In many cases, a small amount of clock skew may be unimportant, especially the skew between distant latches, but performance is expected to suffer from any unmodeled skew. In addition, compared to grid-based clock distributions [10], tree-based networks such as the clock distribution described in this paper consume less power and wiring, but can have more significant unmodeled "local skew" between nearby latches if designed or modeled inaccurately.

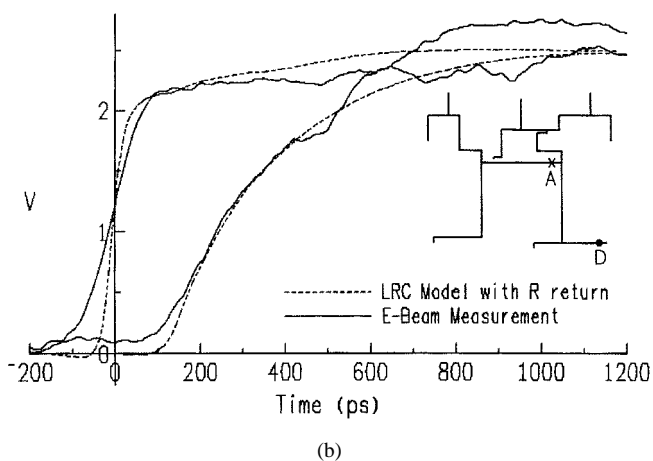
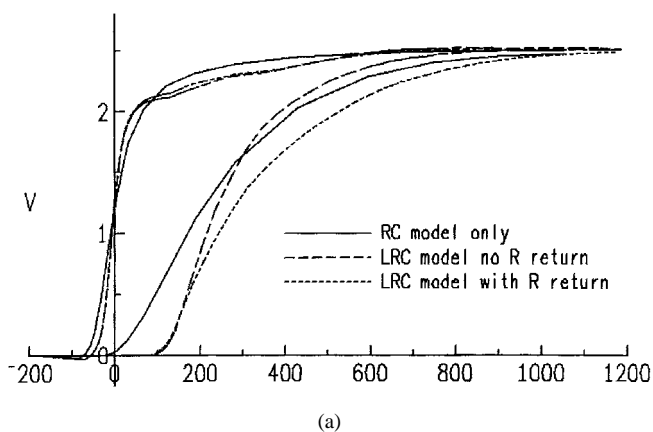


Fig. 5. Test chip waveforms along another branch of the tree, with (a) simulated RC model, LRC model without return path resistance effect, and LRC model with return path resistance effects. (b) Measured waveforms compared to the full LRC simulation. The inset shows the first-level clock tree and the measurement points. The oscillations seen in the measured waves are measurement artifacts.

V. CONCLUSION

The design and modeling of on-chip wiring is becoming more complex and more important for the design of high-performance CMOS microprocessors. Complex lossy trans-

mission line effects have become very significant, but it is possible to obtain excellent results if these effects are considered throughout the design process. Measurement of the clock signals played an essential role in validating the model predictions.

ACKNOWLEDGMENT

The authors would like to thank T. G. McNamara and W. D. Becker of IBM Systems 390 Division and G. V. Kopcsay of IBM Research for design and modeling contributions, and L. T. Pileggi of Carnegie-Mellon University for discussions and communications.

REFERENCES

- [1] A. Deutsch *et al.*, "Modeling and characterization of long on-chip interconnections for high-performance microprocessors," *IBM J. Res. Develop.*, vol. 39, pp. 547–567, Sept. 1995.
- [2] N. Vasseghi, K. Yeager, E. Sarto, and M. Seddighnewzhad, "200 MHz superscaler RISC microprocessor," *IEEE J. Solid-State Circuits*, vol. 31, pp. 1675–1686, Nov. 1996.
- [3] B. A. Gieseke *et al.*, "A 600 MHz superscaler RISC microprocessor with out-of-order execution," in *Digest Tech. Papers 1997 Int. Solid-State Circuits Conf.*, pp. 176–177.
- [4] C. F. Webb, C. J. Anderson, L. Sigal, K. L. Shepard, J. S. Liptay, J. D. Warnock, B. Curry, B. W. Krumm, M. D. Mayo, P. J. Camporese, E. M. Schwartz, M. S. Farrell, P. J. Restle, R. M. Averill III, T. J. Slegel, W. V. Huott, Y. H. Chan, B. Wile, P. G. Emma, D. K. Beece, C.-T. Chuang, and C. Price, "A 400 MHz S/390 microprocessor," *IEEE J. Solid-State Circuits*, vol. 32, pp. 1665–1675, Nov. 1997.
- [5] A. Deutsch *et al.*, "Design guidelines for short, medium, and long on-chip interconnections," in *Proc. 5th IEEE Topical Meet. Electrical Performance of Electronic Packaging*, 1996, vol. EPEP'96, pp. 30–32.
- [6] B. J. Rubin and S. Daijavad, "Calculations of multi-port parameters of electronic packages using general purpose electromagnetics code," in *Proc. IEEE 2nd Topical Meet. Electrical Performance of Electronic Packaging*, 1993, vol. EPEP'93, pp. 37–39.
- [7] C. Ratzlaff and L. T. Pileggi (now Pileggi), "RICE: Rapid interconnect circuit evaluation using asymptotic waveform evaluation," *IEEE Trans. Computer-Aided Design*, vol. 13, pp. 763–776, June 1994.
- [8] Advanced Statistical Analysis Program (ASTAP), "Program Reference Manual," Order no. SH20-118-0; available through IBM branch offices.
- [9] A. J. Fixl and K. A. Jenkins, "Laser stimulated photocathode electron beam prober for 15 ps resolution waveform measurements," *Microelect. Eng.*, vol. 24, pp. 81–88, 1994.
- [10] W. J. Bowhill *et al.*, "Circuit implementation of a 300-MHz 64-bit second-generation CMOS alpha CPU," *Digital Tech. J.*, vol. 7, pp. 100–118, 1997.