Interconnect Modeling and Analysis in the Nanometer Era: Cu and Beyond

Kaustav Banerjee[1], Sungjun Im[2] and Navin Srivastava[1]

[1]Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106, U.S.A.

[2]Department of Materials Science and Engineering, Stanford University, Stanford, CA 94305, U.S.A.

ABSTRACT

This paper highlights key emerging issues in the domain of interconnect modeling and analysis. The implications of various nanoscale effects on VLSI interconnect performance, reliability, power dissipation and parasitic extraction are also presented. Finally, promising new technologies are outlined which have the potential to meet these interconnect challenges in the nanometer era.

INTRODUCTION

As VLSI technology has progressed to pack smaller, faster and increasing number of transistors (doubling every 1.8 years as per Moore's Law) on a single chip than ever before, the demands from wires that connect these devices have increased tremendously. In fact, Copper/Low-k interconnect technologies for sub-100 nm CMOS ICs are impacting system performance through increased power dissipation, signal delay, and cross-talk. With clock frequencies increasing into the GHz regime, the parasitic resistance, capacitance and inductance associated with these wires often lead to performance bottlenecks which have led the semiconductor and the electronic design automation (EDA) industries to adopt several technological innovations. Furthermore, prevalent high chip temperatures (aggravated by large power dissipation of nanometer scale ICs) and increasing current densities in wires make electromigration in copper a constant threat to VLSI circuits. Moreover, variability issues at the nanometer scale are evident in interconnects as they are in devices. Thus, interconnects have come to share the center-stage alongside logic devices in terms of design and analysis efforts. This paper highlights key emerging issues in the domain of interconnect modeling and analysis, their implications as well as some promising new technologies that have the potential to meet these interconnect challenges in the nanometer era.

I. Material Issues

1. Copper resistivity increase at nanometer scale metal dimensions

With the aggressive scaling of VLSI technology, cross-sectional dimensions of on-chip interconnects in current technologies [1] are of the order of the mean free path of electrons in copper (40 nm at room temperature). At such dimensions, the increase in Cu interconnect resistivity due to the presence of a highly resistive barrier layer (which occupies a significant fraction of the drawn wire width) is further exacerbated by the increased scattering of electrons at the surface and grain boundaries [2]. Fig. 1(a) shows the effective resistivity of Cu interconnects, along with the contributions of different factors for future technology nodes [3]. A discussion of the models used for these calculations can be found in [2]. It can be observed that the contribution of surface scattering and grain boundary scattering is roughly the same, but both increases with scaling. The background scattering of the electrons by phonons, electrons and defects (impurities) contributes to the bulk resistivity value (ρ_o) that remains constant.

2. Thermal conductivity of low-k dielectrics

The low-k dielectric materials used for intra- and inter- metal layer dielectric (ILD) in current VLSI technologies have much lower thermal conductivities than silicon dioxide. Fig. 1(b) compares predictions based on the differential-effective-medium (DEM) and the porosity weighted simple medium (PWSM) models for porous silicate (Xerogel) films with experimental data for various dielectrics from various sources in the published literature [3].

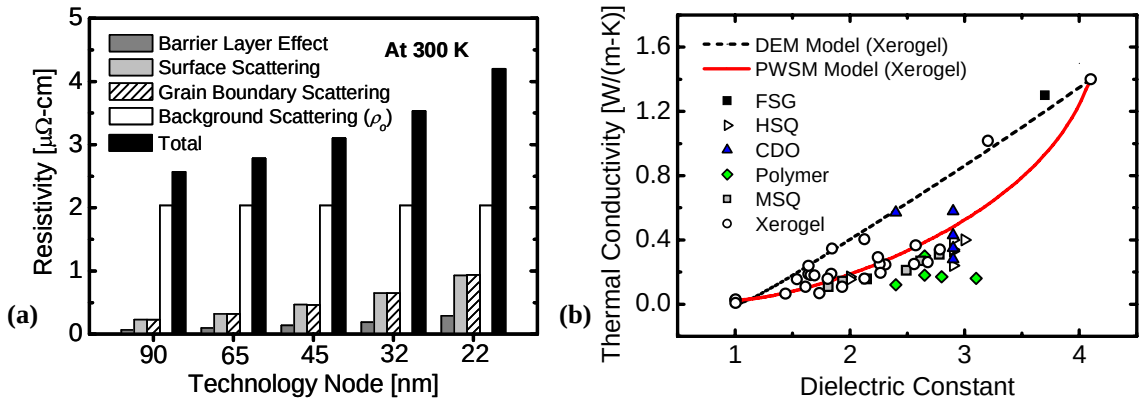

Fig. 1: (a) Scaling of metal resistivity for the ITRS intermediate tier wires (at 300K) [3]. **(b)** Correlation of the thermal conductivity and dielectric constant of ILDs. Inter-layer dielectric constant is lowered as technology scales (for improvement in performance), but thermal conductivity of the dielectric also decreases rapidly [3].

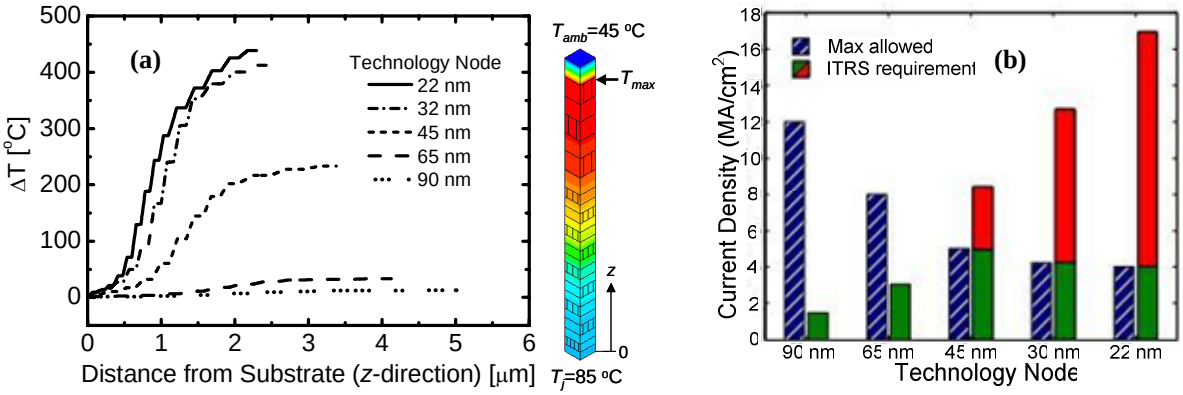

Fig. 2. (a) Temperature rise (ΔT) along the vertical distance from the substrate at different technology nodes [3]. The temperature contour plot of 45 nm technology node is also shown as an example. **(b)** Maximum allowed current density (duty ratio=0.001) in local vias from self-consistent electromigration lifetime estimation [5] vs. the ITRS requirement for current density.

3. Implications

a. Implications for Temperature and Reliability

Alongside the lower thermal conductivity of inter-layer dielectrics, the scaling of technology also results in higher current density demands on interconnects [4]. The combined effect of increasing copper interconnect resistivity, decreasing thermal conductivity of ILD materials and rising current densities in on-chip wires results in significant rise in interconnect temperatures, especially at the global metal layers which are furthest away from the heat sink (Fig. 2(a)). These high temperatures will become a major concern for interconnect reliability as the mean time to failure due to electromigration depends exponentially on metal temperature. The maximum current density that can be supported by these interconnects will thus be severely limited due to reliability constraints, as shown in Fig. 2(b) [4].

b. Implications for Latency and Power Dissipation

While typical local interconnect delay is expected to decrease with technology scaling (mainly as a result of the higher packing density of devices), global interconnect delay increases as seen in Fig. 3(a). In order to keep the delay of global interconnects under control, repeaters (inverters) are inserted at regular intervals to drive signals faster. As technology scales, an increasing number of buffers is required for the global interconnection system on a chip (Fig. 3(b)). These repeaters can contribute significantly to total chip power dissipation, which is a critical problem for high-performance ICs.

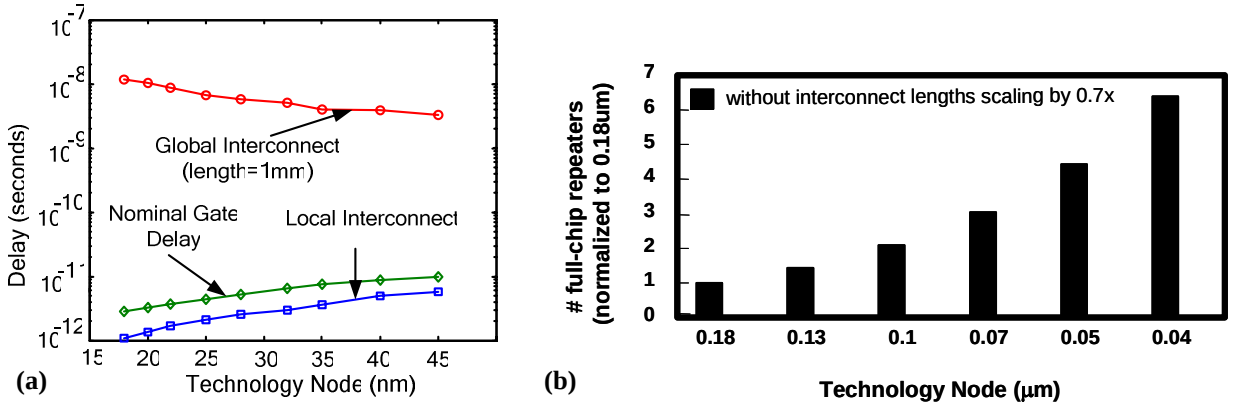

Fig. 3: (a) Typical delay for global and local Cu interconnects compared to nominal gate delay [1], as technology scales. **(b)** Number of repeaters increases rapidly with technology scaling [courtesy of Intel].

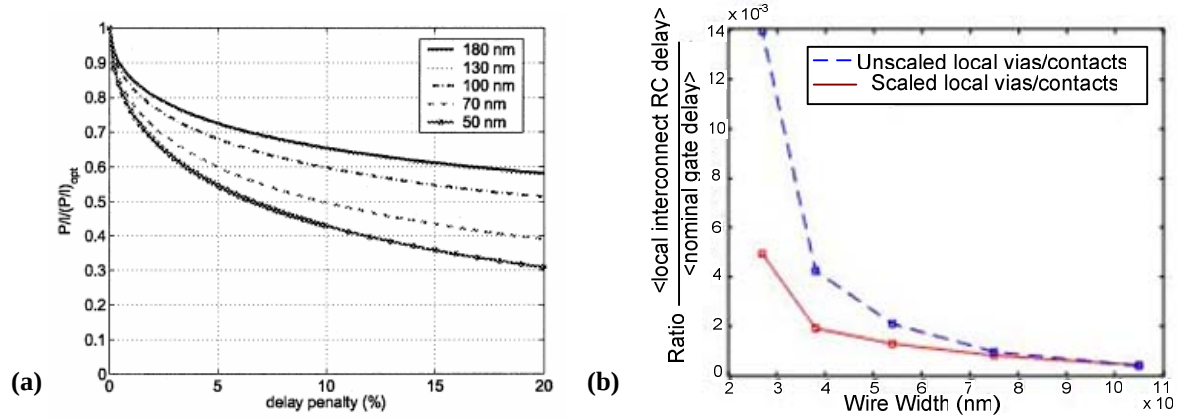

Fig. 4: (a) Normalized power per unit length for a buffered global interconnect (normalized to power per unit length for optimally buffered case) as a function of delay penalty [6]. **(b)** Ratio of local interconnect RC delay (including 2 contacts + 2 local vias) to nominal gate delay as per ITRS, as technology scales [4].

Increase in effective wire resistivity will further increase wire delay resulting in more numerous repeaters and larger chip power dissipation. However, it has been shown that when delay is not of critical importance "power-optimal" repeater insertion [6] can be used to achieve large power savings. This methodology assumes tremendous importance in the light of power-limited technologies of current and future IC generations as the power savings for a given amount of delay penalty increases as technology scales (Fig. 4(a)). This is mainly due to the increasing leakage power dissipation in CMOS devices. Moreover, the traditional belief about the decrease in local interconnect delay as a result of technology scaling also does not remain true for scaled technologies, mainly as a result of the steep increase in resistivity of small dimension local vias and contacts, which have the smallest dimensions among all wires on a chip. This effect is evident in Fig. 4 (b). It is also evident that the increase in local interconnect delay with scaling is significantly steeper if via (and contact) heights are not scaled.

II. Electrical Issues

Any transient signal carried by an interconnect on a chip, for example, a rising clock edge, is composed of numerous constituent signals spread across the frequency spectrum. The significant frequency - the highest frequency up to which the amplitude of the corresponding frequency components is significant – which depends on the signal rise time, is much higher than the frequency of the clock in a digital system as shown in Fig. 5(a). At these frequencies, it becomes important to consider inductive effects especially in the unscaled global interconnects [7]. Moreover, the skin depth (penetration depth of electromagnetic waves) for copper, at these high frequencies, is smaller than the typical cross-sectional dimensions of global wires. This, known as the “Skin Effect”, can increase the effective resistance of wires.

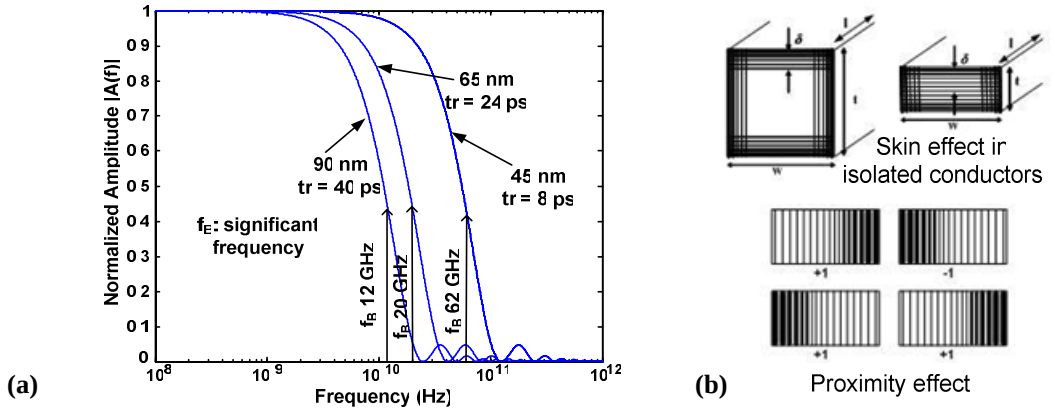

Fig. 5: (a) Frequency components of a digital VLSI signal at different technology nodes. The signal rise time (obtained from SPICE simulations) and significant frequency at each technology node are also shown. **(b)** Schematic depicting current density distribution under skin and proximity effects.

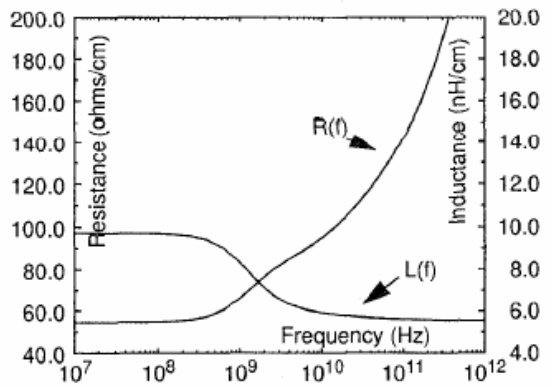

Fig. 6: Frequency dependence of interconnect resistance and inductance as a result of skin and proximity effects [8].

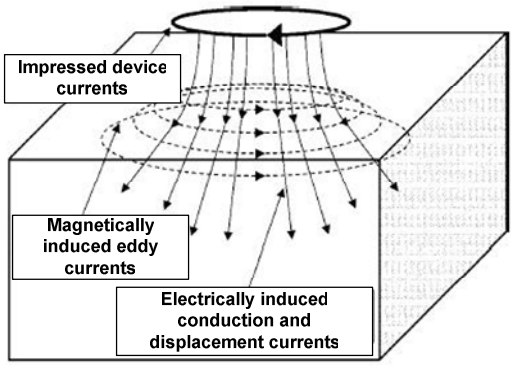

Fig. 7: Schematic representation of magnetically induced substrate eddy currents [11].

Besides the Skin Effect, the current distribution within the conductors is also altered due to the presence of proximal current carrying conductors at high frequencies (proximity effect), as shown schematically in Fig 5(b). Fig. 6 shows the impact of these high frequency effects (Skin and Proximity) on the impedance of a typical interconnect structure as a function of frequency [8]. Since there are no analytical formulae to calculate interconnect impedance under these high frequency effects, the accuracy of computationally expensive electromagnetic field solvers such as [9] used for extracting interconnect parasitics (RLC), has to be traded-off against the speed but limited accuracy of equivalent circuit models such as those presented in [8, 10]. Another complexity in the extraction of interconnect impedance arises from the effect of "Eddy Currents" that are magnetically induced in the finite conductivity substrate (depicted schematically in Fig. 7). Substrate effects have a significant impact on interconnect impedance [12] and especially on RF components such as on-chip inductors [13]. The analysis of long range coupling crosstalk between interconnects also becomes increasingly challenging in the presence of a lossy substrate [14]. These substrate effects add yet another dimension to the complexity of interconnect parasitic extraction [11].

III. Variability Issues

As shown in Fig. 8(a), the gap between polysilicon gate lengths and the wavelength of light used in optical lithographic process is increasing with technology scaling, as a result of which, MOSFET device dimensions show significant amount of variations [15]. As local interconnect widths closely follow the scaling of gate lengths at any technology node, interconnect dimensions also suffer similar variations.

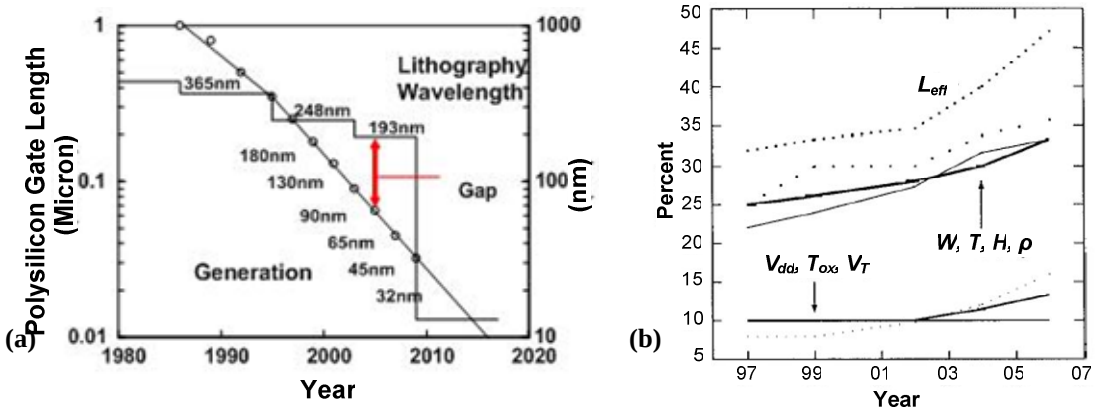

Fig. 8: (**a**) Increasing gap between polysilicon gate length and lithographic wavelength for different technology nodes [15]. (**b**) Percent variations in device (V_{dd}, T_{ox}, V_T and L_{eff}) and interconnect (W, T, H, ρ) parameters as technology scales [16].

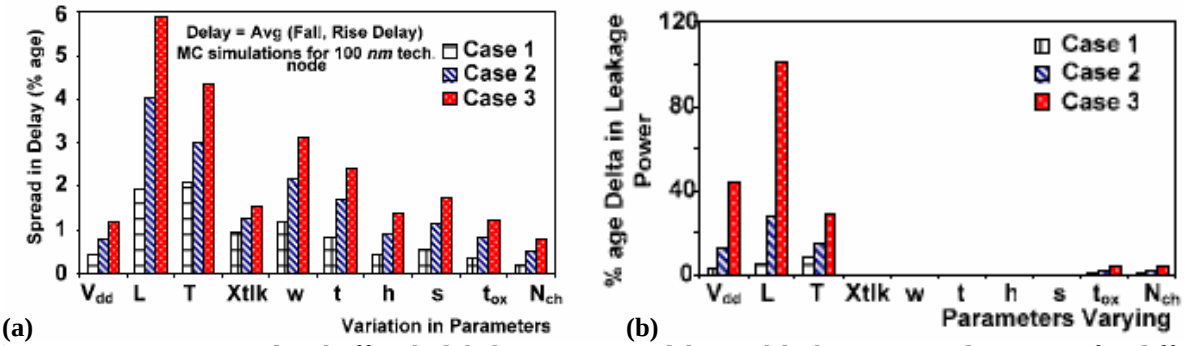

Fig. 9: Percentage spread in buffered global interconnect delay and leakage power dissipation for different percentage variations (Case 1: 10%, Case 2: 20%, Case 3: 30%) in different parameters (V_{dd}: supply voltage, L: channel length, T: temperature, Xtlk: interconnect crosstalk, w: interconnect width, t: interconnect thickness, h: inter-layer dielectric thickness, s: wire spacing, t_{ox}: gate oxide thickness and N_{ch}: doping concentration) [19].

In addition to the lack of control on extremely scaled wire widths at the local interconnect level, the wire height in wide global interconnects also undergoes significant variations as a result of dishing of metal during the chemo-mechanical polishing (CMP) process which follows metal deposition in each metal layer (Fig. 8(b)).

Variations in interconnects need to be taken into account for the purpose of parasitic extraction as interconnect geometry variations can have significant impact on circuit timing [16, 17]. The work in [18] presents a methodology to extract interconnect parasitic capacitance in the presence of process variations. The work in [19] analyzes the impact of variations in devices as well as interconnects on the global interconnect buffer insertion problem using a statistical framework. It is shown that leakage power in a buffered interconnect system is predominantly sensitive to device variations, whereas delay is sensitive to both device and interconnect variations (Fig. 9). It is also shown that the benefit derived from "power-optimal" repeater insertion increases when variations are taken into account.

IV. Emerging Interconnect Technologies

The challenges from on-chip interconnects in nanometer scale VLSI have led researchers to seek innovative design solutions, circuit or interconnect optimization techniques and material solutions so that the chip's wires do not offset the benefits of continued device scaling. Three-dimensional integration to create multi-layer Si chips [20] is a concept that reduces the number and the average lengths of the longest global wires seen in traditional 2-D chips by providing shorter "vertical" paths for connection. However, this technology still needs to overcome difficult challenges such as the thermal management of

internal (stacked) active layers, development of new system architectures and tools, etc. Optical interconnects [21] and radio frequency or wireless interconnects [22] are also being considered as alternatives to traditional wires for providing chip-to-chip as well as intra-chip communication. Their major benefits are the high propagation speeds and the fact that they are not bandwidth limited like electrical interconnects. However, their choice as a replacement technology is limited by the need for specialized components such as transmitter/receiver circuits and error correction circuits which must be placed wherever such interconnects are used.

Carbon Nanotube Interconnects

Carbon nanotube (CNT) interconnects are possibly the least disruptive of all alternatives to copper interconnects that have been suggested so far. Although there are some technological issues that must be resolved before CNT interconnects can be used in practice, they have the potential to meet interconnect challenges without the need for paradigm changes in VLSI circuit design techniques and tools (as in the case of 3-D ICs) or the need for extra circuitry (as in the case of optical or RF interconnects).

CNTs are graphene sheets rolled up into cylinders (Fig. 10(a)) [23] with diameter of the order of a nanometer. Because of their extremely desirable properties of high mechanical and thermal stability, high thermal conductivity and large current carrying capacity [24], CNTs have aroused a lot of research interest in their applicability as VLSI interconnects of the future. Bundles of metallic single-walled CNTs (SWCNTs) with electron mean free paths of the order of a micron are the most suitable candidates for interconnects. Fig. 10(b-d) shows different strategies reported in the literature that can potentially be used for forming interconnects and vias using carbon nanotube bundles. The performance comparison between Cu interconnects and interconnects formed from bundles of carbon nanotubes, based on the performance model described in [27], is summarized in Fig. 11.

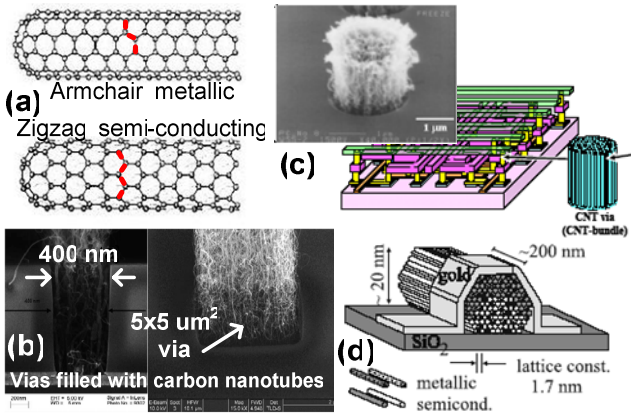

Figure 10: (a) Different configurations and resulting electrical conduction types of carbon nanotubes depending on the direction along which the graphene sheets are rolled up (chirality) [23]. **(b-d)** Strategies to form interconnects and vias using carbon nanotube bundles (b: [24], c: [25], d: [26]).

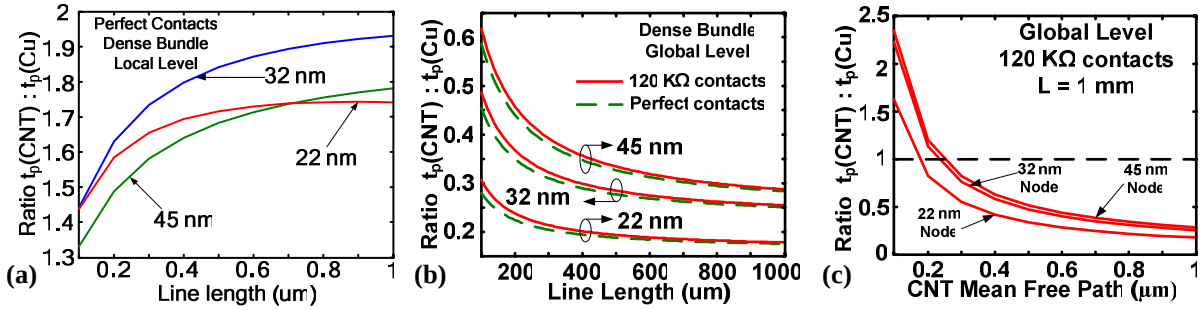

Figure 11: Ratio of interconnect propagation delay with CNT bundle interconnect to that with Cu wire of same dimensions for **(a)** local interconnect and **(b)** global interconnect, as a function of interconnect length, and **(c)** for global interconnect as a function of CNT mean free path [27].

At the local interconnect level, CNT-bundles do not give much performance improvement. However, in the case of long intermediate and global interconnects, densely packed CNT bundle interconnects show significant improvement in performance as compared to copper interconnects, even with imperfect metal-nanotube contacts. Hence, if long lengths of densely packed carbon nanotube bundles with mean free path around 1um can be fabricated reliably, they will provide a valuable alternative to copper interconnections. Besides the performance benefits of CNT bundle interconnects, their high thermal conductivity makes them very effective in controlling the large backend temperature rise expected with metallic interconnects (Fig. 2(a)), and hence in improving overall interconnect performance and lifetime [28].

SUMMARY

The increasing copper resistivity and low thermal conductivity of inter- and intra- layer dielectric materials are identified as key interconnect technology issues that can threaten the performance as well as reliability of sub-45 nm copper interconnects. The electrical issues mainly as a result of high frequency signals make interconnect parasitic extraction and analysis a daunting task which is compounded because of variability in interconnects. Carbon nanotube interconnects are identified as promising candidates to meet these challenges faced by metallic interconnects.

REFERENCES

1. International Technology Roadmap for Semiconductors (ITRS), 2004, http://public.itrs.net
2. W. Steinhogl, et al., *J. App. Physics*, Vol. 97, No. 2, pp. 023706-1-023706-7, 2005.
3. S. Im, N. Srivastava, K. Banerjee and K. E. Goodson, *IEEE Trans. on Electron Devices*, 2005 (to appear).
4. N. Srivastava and K. Banerjee, *Proc. 21st Intl. VLSI Multilevel Int. Conf.*, 2004, pp. 393-398.
5. K. Banerjee and A. Mehrotra, *IEEE Circuits and Devices Magazine*, pp. 16-32, 2001.
6. K. Banerjee and A. Mehrotra, *IEEE Trans. on Electron Devices*, Vol. 49, No. 11, pp. 2001-2007, 2002.
7. K. Banerjee and A. Mehrotra, *IEEE Trans. on CAD*, Vol. 21, No. 8, pp. 904-915, 2002.
8. B. Krauter and S. Mehrotra, *Design Automation Conference*, 1998, pp. 303-308.
9. Z. Zhu, B. Song and J. White, *Design Automation Conference*, 2003, pp. 712-717.
10. S.-P. Sim, et al, *IEEE Trans. on Electron Devices*, Vol. 50, No. 6, pp. 1501-1510, 2003.
11. A. M. Niknejad and R. G. Meyer, *IEEE Trans. on Microwave Theory and Techniques*, Vol. 49, No. 1, pp. 166-176, 2001.
12. D. De Roest, et al, *IEEE Electron Device Letters*, Vol. 23, No. 2, pp. 103-104, 2002.
13. A. Pun, et al, *Intl. Electron Devices Meeting*, 1997, pp. 325-328.
14. Y. Massoud and J. White, *IEEE Trans. on VLSI Systems*, Vol. 10, No. 3, pp. 286-291, 2002.
15. P. Gelsinger, 41st DAC Keynote, *Design Automation Conference*, 2004, www.dac.com
16. D. Boning and S. Nassif, in *Design of High Performance Microprocessor Circuits*, Edited by Chandrakasan, Bowhill and Fox, Wiley-IEEE Press, 2000.
17. Z. Lin, et al, *IEEE Trans. on Semiconductor Manufacturing*, Vol. 11, No. 4, pp. 557-568, 1998.
18. A. Labun, *IEEE Trans. on CAD*, Vol. 23, No. 6, 2004.
19. V. Wason and K. Banerjee, *Intl. Symposium on Low-Power Electronic Design*, 2005, pp. 131-136.
20. K. Banerjee, et al, *Proceedings of the IEEE*, Vol. 89, Issue 5, pp. 602-633, 2002.
21. D. A. B. Miller, *Proceedings of the IEEE*, Vol. 88, Issue 6, pp. 728-749, 2000.
22. B. Floyd, K. Kim and O. Kenneth, *IEEE Intl. Solid-State Circuits Conference*, 2000, pp. 328-329.
23. P. L. McEuen, M. S. Fuhrer and H. Park, *IEEE Trans. Nanotechnology*, Vol. 1, No. 1, pp. 78-85, 2002.
24. F. Kreupl, et al., *Microelectronic Engineering*, 64 (2002), pp. 399-408.
25. M. Nihei, M. Horibe, A. Kawabata and Y. Awano, *IEEE Intl. Int. Tech. Conf.*, 2004, pp. 251-253.
26. H. Stahl, et al., *Physical Review Letters*, Vol. 85, No. 24, pp. 5186-5189, 2000.
27. N. Srivastava and K. Banerjee, *IEEE Intl. Conf. on Computer-Aided Design*, 2005 (to appear).
28. N. Srivastava, R. V. Joshi and K. Banerjee, *IEEE Intl. Electron Devices Meeting*, 2005 (to appear).