Interconnect Challenges for Nanoscale Electronic Circuits

Navin Srivastava and Kaustav Banerjee

This article provides an overview of the new challenges by nanometer-scale on-chip interconnects. Effects on performance and reliability are addressed, with an emphasis on resistivity, interconnect delay, and current-carrying capability.

INTRODUCTION

Transistors operate faster as their dimensions are scaled down. The wires on the chips that connect these transistors to form a circuit, however, do not exhibit the same benefit of scaling. The drive for faster chips with lower cost and greater functionality has transformed these wires (interconnects) into what determines the performance and reliability of a nanometer-scale integrated circuit (IC). This paper provides an overview of the nanometer-scale interconnect issues beyond apparent technology scaling effects, with an emphasis on challenges beyond 90 nm.

NANOMETER-SCALE EFFECTS ON INTERCONNECTS

Resistivity

Copper has recently replaced aluminum as the dominant interconnect material. A drawback with copper, however, is the need for a diffusion barrier to prevent copper from diffusing into the surrounding dielectric. This film has much higher resistivity than copper, and ~20% of the wire width can be consumed by the barrier film. In addition, wire cross-sectional dimensions are on the order of the mean free path of electrons (~40 nm at room temperature).[1] At such dimensions, the electron-scattering effect at the conductor surface as well as at the grain boundaries cause its resistance to increase. Copper interconnect resistivity increases by several times over its bulk value (1.9 micro-ohm per centimeter at room temperature)[1] at sub-90 nm technology nodes (Figure 1.)

Interconnect Delay

The intrinsic delay incurred on short-length wires has remained low compared to the logic gate delay. However, with the size effect mentioned previously, the local interconnect delay starts to increase significantly. Figure 2 compares the intrinsic local interconnect (traversing two contact plugs and two vias) delay with the logic delay. The local interconnect delay shows a sharp increase as the technology progresses beyond the 45 nm node, though it remains less prevailing than the gate delay. At the global level, however, the interconnect delay continues to dominate the logic delay as global signals traverse long distances across the chip. With technology scaling and added functionality, the number and length of these global lines increases. Since the delay of a long unbuffered line is quadratic in its length, long wires are divided into smaller segments using repeaters or buffers. The delay of an optimally buffered line is linear in its length.[2] Hence, repeater insertion is an effective way to keep global wiring delay under control in IC design. Furthermore, the integration of low dielectric-constant materials can reduce interconnect capacitance and delay. Even with these techniques, however, the global signal delay constitutes one of the biggest challenges for future interconnects.[3]

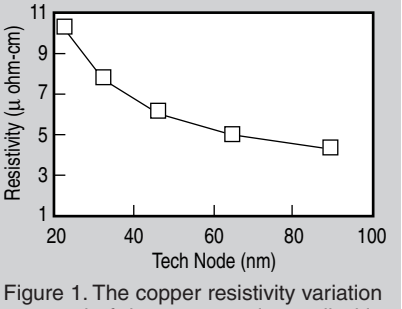

Figure 1. The copper resistivity variation as a result of electron scattering applicable to nanoscale interconnects (T = 120°C).

For large high-performance designs, the number of repeaters can be prohibitively high[4] (in excess of 10^6 for sub-90 nm designs). In general, the repeaters are optimally sized and separated. However, since these repeaters are large, the total power dissipation by such repeaters can be unacceptably large. The power dissipation is exacerbated by the substantial rise in leakage power beyond sub-90 nm technologies. It has been shown[4] that by incurring a small delay penalty on global signal lines that do not lie on the critical timing path, a potential solution exists for large power savings by using smaller and fewer repeaters. The optimization scheme[4] becomes important in nanometer-scale technologies where reduced power dissipation is a key performance criterion.

Current-Carrying Capability: Reliability

The current-carrying capacity of interconnects is limited by Joule heating and electromigration (EM). Joule heating is affected by the resistance of the conducting wire. For contacts and local vias that have the smallest cross-sectional dimensions among on-chip interconnects, current-carrying capability is a concern. The current density required to be carried by local vias and contacts

needs to increase at a much faster rate than that for other interconnects as technology scales beyond 90 nm.

The maximum allowable current density for interconnects is also dependent on EM lifetime, which is exponentially dependent on the metal temperature. Figure 3 shows that EM and thermal constraints make the current density requirements unachievable beyond 45 nm technology node. (Arrows along the y-axis show current density needed to support International Technology Roadmap for Semiconductors (ITRS) requirements.) The metal temperature, in turn, is dependent on the junction (die) temperature of the chip. Due to threshold voltage scaling and process variations, leakage power dissipation has become a significant component of total chip power dissipation in nanometer scale-IC designs. Since the sub-threshold leakage current is exponentially dependent on temperature, the total chip power dissipation also becomes a function of temperature. This marks a significant departure from the traditional method of junction temperature evaluation. A complete methodology has been developed for the accurate computation of junction temperature using the electro-thermal couplings between temperature, power, voltage, and frequency.[5] These electrothermal couplings have direct implications on the metal temperature, and hence on the EM lifetime.

EMERGING INTERCONNECT TECHNOLOGIES

The challenges in nanoscale interconnects have led researchers to seek

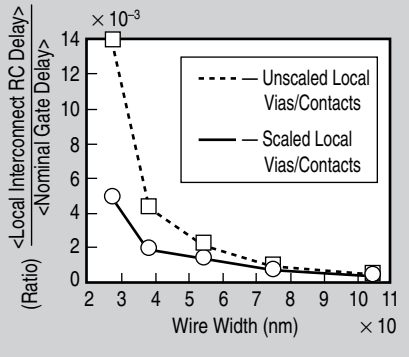

Figure 2. The ratio of local interconnect RC delay to nominal gate delay as technology scales.

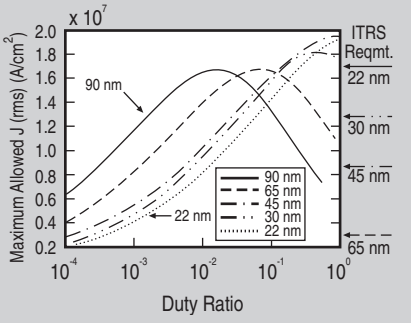

Figure 3. Maximum allowed RMS current densities for local local vias under self-consistent Joule heating and EM lifetime constraints.

innovative design and/or technological alternatives. Three such approaches are outlined in the following.

3-D ICs

Three-dimensional (3-D) integration to create multi-layer silicon chips promises to alleviate the challenge of global interconnects in nanometer-scale circuits.[6] A 3-D integration scheme reduces the number and the average length of the longest global wires in two-dimensional chips by providing shorter vertical paths for connection. Besides the interconnect performance benefits, 3-D integration leads to increased transistor packing density and smaller chip area and provides means to integrate dissimilar technologies (digital, analog, radio frequency circuits, etc.) in the same chip but on different active layers. However, this technology needs to overcome difficult challenges such as thermal management and development of new system architectures and design tools.

Optical Interconnects

Optical interconnects, which exploit the propagation speed of light, are considered an attractive alternative for providing chip-to-chip as well as intra-chip communication.[7] They use either guided waves or free-space optics with conventional lens or microlens arrays to transmit signals on long communication lines. The major benefits of optical interconnects are the high propagation speeds and the fact that they are not bandwidth limited. However, their practicality is limited by the need for specialized opto-electronic components and error-correction circuits that must be placed wherever optical interconnects are used.

Carbon Nanotubes

Carbon nanotubes (CNTs) exhibit high current-carrying capacity, mechanical stability. and thermal conductivity. Numerous applications, including interconnects, have been proposed for CNTs.[8] They have diameters of the order of a nanometer and can sustain high current stress of ~1×10^9 A/cm^2 [9] without failing due to electromigration. The resistance of a single ballistic single-walled CNT (SWCNT) less than 1 μm long, assuming perfect contacts, is about 6.45 k-ohms.[10] This is the fundamental resistance associated with an SWCNT. The resistance of a CNT necessitates the use of an array of a number of parallel CNTs in order to realize a low-resistance interconnect. Carbon nanotube arrays are a potential alternative to metallic contacts and vias at the local level, provided they can be grown with sufficiently high densities.[11]

References

1. W. Steinhogl et al., *Phys. Rev. B*, 66 (2002) 075414.
2. H.B. Bakoglu, *Circuits, Interconnections and Packaging for VLSI* (Reading, MA: Addison-Wesley, 1990).
3. "International Technology Roadmap for Semiconductors" (2003), *public.itrs.net.*
4. K. Banerjee and A. Mehrotra, *IEEE Trans. Electron Dev.*, 49 (11) (2002), pp. 2001–2007.
5. K. Banerjee et al., *49th Annual IEEE International Electron Devices Meeting*, (Piscataway, NJ: IEEE, 2003) pp. 887–890.
6. K. Banerjee et al., *Proc. IEEE*, 89 (5) (2001), pp. 602–633.
7. D.A.B. Miller, *Proc. IEEE*, 88 (6) (2000), pp. 728–749.
8. F. Kreupl et al., *Microelectronic Engineering*, 64 (2002), pp. 399–408.
9. B.Q. Wei, R. Vajtai, and P.M. Ajayan, *Appl. Phys. Lett.,* 79 (8) (2001), pp. 1172–1174.
10. P.L. McEuen, M.S. Fuhrer, and Hongkun Park, *IEEE Trans. Nanotech.,* 1 (1) (2002), pp. 78–85.
11. N.Srivastava and K. Banerjee, *Proc. 21st Intl. VLSI Multilevel Interconnect Conf., (*Tampa, FL: IMIC, 2004).

Navin Srivastava and Kaustav Banerjee are with the Department of Electrical and Computer Engineering at the University of California at Santa Barbara in Santa Barbara, California.

For more information, contact Kaustav Banerjee, University of California at Santa Barbara, Department of Electrical and Computer Engineering, Room 4151, Engineering I, Santa Barbara, CA; e-mail kaustav@ece.ucsb.edu.