On the Applicability of Single-Walled Carbon Nanotubes as VLSI Interconnects

Navin Srivastava, Student Member, IEEE, Hong Li, Student Member, IEEE, Franz Kreuapl, Member, IEEE, and Kaustav Banerjee, Senior Member, IEEE

Abstract—This paper presents a comprehensive study of the applicability of single-walled carbon nanotubes (SWCNTs) as interconnects in nanoscale integrated circuits. A detailed analysis of SWCNT interconnect resistance (considering its dependence on all physical parameters, as well as factors affecting the contact resistance), the first full 3-D capacitance simulations of SWCNT bundles for realistic very large scale integration (VLSI) interconnect dimensions, and a quantitative evaluation of the importance of inductive effects in SWCNT interconnects are presented. The applicability of carbon nanotube (CNT) based vias (vertical interconnects)—the most realizable CNT interconnects in the current state of the art—is addressed for the first time. It is shown that CNT interconnects can provide 30%–40% improvement in the delay of millimeter-long global interconnects. The applicability of CNT monolayers as local interconnects is found to be much more limited than that reported in the prior literature. Dense CNT bundle global interconnects are shown to offer a 4 × reduction in power dissipation while achieving the same delay as optimally buffered Cu interconnects at the 22 nm node. This power saving increases to 8 × at the 14 nm node. Furthermore, 3-D finite-element electrothermal simulations show that CNT bundles used as vias in between Cu metal layers can provide large improvement in metal interconnect lifetime by lowering the temperature of the hottest interconnects.

Index Terms—Carbon nanotube (CNT), performance, power, reliability, thermal management, very large scale integration (VLSI) interconnect.

I. INTRODUCTION

SINCE their accidental discovery in 1991 [1], carbon nanotubes (CNTs) have aroused a lot of interest as building blocks for future integrated circuits (ICs) due to their outstanding electrical properties [2], high thermal and mechanical stability, high thermal conductivity, and large current-carrying capacity [3]–[6]. Alongside research efforts into developing semiconducting CNT-based transistors [7], metallic CNTs (mCNT) have also been suggested as an interconnect material [8], [9]. Several groups have reported progress on the fabrication of CNT-based interconnects [8], [10], [11], which can alleviate the significant performance and electromigration (EM) reliability concerns in future generations of Cu interconnects [12], [13]. The widening gap between ITRS current-density requirements [14] and the current-carrying capacity of Cu vias—primarily due to the steep rise in parasitic resistance of aggressively scaled Cu interconnects [14]–[16] and rising on-chip metal temperatures—is shown in Fig. 1, indicating the need for identifying alternate interconnect materials for very large scale integration (VLSI).

In this paper, we present a comprehensive study of the applicability of single-walled CNTs (SWCNTs) as VLSI interconnects. The paper is organized as follows. Section II provides a brief introduction to CNT interconnect technology. Section III starts with a review of existing works analyzing the performance of CNT interconnects and presents an equivalent circuit model, based on the Luttinger liquid theory, to calculate the propagation delay of a CNT bundle interconnect. The properties of the equivalent circuit model are discussed, highlighting all underlying assumptions. Salient features of Section III include the impact of all physical parameters on CNT resistance, the first 3-D capacitance simulation results for CNT bundles of realistic VLSI interconnect dimension, and a quantitative analysis of the importance of inductance in calculating CNT interconnect propagation delay.

Based on the equivalent circuit model, the performance of CNT bundles is evaluated in Section IV, and the advantages...
they can offer at different domains of on-chip interconnects are shown. This performance analysis provides a clear insight into the applicability of CNT interconnects in the form of dense or sparse bundles as well as monolayers. Section IV also addresses the impact of CNT bundle interconnects on global interconnect power dissipation—a critical concern for future technology generations and various interconnect-centric microarchitectures. Noting recent progress in the fabrication of CNT vias, Section V addresses the concept of hybrid CNT/Cu interconnects [9]—employing CNT vias in tandem with Cu metal layers—and its remarkable advantages from a reliability/thermal-management perspective. This paper is unique in its key findings with regard to interconnect power dissipation and thermal analysis of the back-end interconnect stack. Section VI concludes with a summary of our findings and lays out the process technology requirements that must be met in order to make SWCNT interconnects a reality.

II. CNT INTERCONNECTS

The basic structure of a CNT—as a rolled-up graphene sheet to form a SWCNT, or several concentric tubes to form a multiwalled CNT (MWCNT)—is shown in Fig. 2. Changing the direction of roll-up leads to different “chiralities,” which are identified by the chiral indexes \((m, n)\) [17, 18] (when either \(m\) or \(n\) is zero it is called zigzag chirality, \(m = n\) is called armchair, while all other nanotubes are called chiral). SWCNTs are metallic when the difference between the chiral indexes is an integer multiple of 3 \((m - n = 3i)\), and are semiconducting otherwise. Hence, armchair SWCNTs are always metallic. Metallic CNTs have aroused a lot of interest as VLSI interconnects of the future [8]–[11], [19]–[22], because of their extremely desirable electrical and thermal properties. CNTs have shown current-carrying capacity in excess of \(10^{10}\) A/cm\(^2\) without any signs of damage even at an elevated temperature of 250\(^\circ\)C [3, 23], thereby eliminating the EM reliability concerns that plague nanoscale Cu interconnects.

Although the current-carrying capacity of an isolated CNT is much greater than that of Cu, it has a high resistance (theoretical minimum of 6.45 kΩ [24, 25]). In practice, the observed resistance of an individual CNT can be much higher [25, 26], due to the presence of imperfect metal–nanotube contacts. The high resistance of an individual CNT necessitates the use of a bundle (rope) of CNTs conducting current in parallel to form a low-resistance interconnect [8], [10], [11].

The work by Kreupl and coworkers [8], [19] proposed a method for growing MWCNT bundles by chemical vapor deposition (CVD) from the bottom of vias and contact holes using an iron-based catalyst. In an alternative bottom-up process [11], MWCNTs are first grown (using HF-CVD) at prespecified locations, then gap-filled with oxide, and finally planarized. The high temperature (700\(^°\)C) during CNT growth [8] has been lowered to 510\(^°\)C in [19], and more recently to a low of 365\(^°\)C in [27]. Advances have also been reported on the reduction of metal–nanotube bundle contact resistance [22], integration with ultra-low K dielectrics [27], CNT through-silicon-vias [28], and densification techniques for SWCNT [29] and MWCNT [30] bundles.

While MWCNT bundles are easier to fabricate, dense metallic SWCNT bundles have greater advantages in terms of interconnect performance [31] and reliability [32]. SWCNT bundle fabrication has been difficult due to the low fertility of catalyst particles for SWCNT growth. It has recently been shown that the fertility of the growth catalyst can be increased by addition of water [33] or oxygen [34]. Thus, homogeneous SWCNT films have been developed by assembling SWCNTs into a dense “carbon nanotube wafer,” enabling integration of nanodevice systems by lithography [35]. Nevertheless, the fabrication of on-chip SWCNT interconnects still poses several challenges because of the need for low-thermal-budget dense metallic nanotube bundles.

III. MODELING AND ANALYSIS OF SWCNT INTERCONNECTS

As the interest in CNT-based interconnects gains momentum, this paper addresses the need to identify the domains of on-chip interconnections (local, intermediate, and global interconnects and vias), where this technology has the potential to replace Cu. It is shown that the requirements from SWCNT process technology vary significantly with the interconnect application that is targeted. This paper analyzes the performance, power, and reliability of SWCNT bundle interconnects in comparison with Cu, taking into account their advantages as well as practical limitations, and lays out the process requirements to make it a viable interconnect technology.

A. Previous Work

The imminent thermal/reliability issues of Cu vias and contacts and the advantages of CNT bundle vias have been reported in [12].

The analysis in [36] considers a flat array of CNTs [Fig. 3(b)], and reports that CNT interconnects have much larger delay than Cu. A similar analysis in [37] arrives at the same conclusions. However, the comments by Li et al. [38] show that the analysis presented in [36] and [37] has several drawbacks. Many other works in the published literature have shown that CNT-based interconnects compare favorably with Cu. Their contributions and limitations are discussed in the following paragraphs.

The superior performance of CNT bundle interconnects as compared with Cu has been shown in [39] and [40]. However, these works neglect practical limitations such as the low density of metallic nanotubes and the presence of imperfect metal–nanotube contacts. Moreover, it is assumed that CNT bundles always have the same capacitance as Cu wires. The work in [41]...
Fig. 3. (a) Equivalent circuit model for an isolated SWCNT [51], for length less than the mean free path of electrons. (b) CNT arrays and (c) CNT bundles increase the number of nanotubes to reduce effective resistance ($n_x$ and $n_y$ denote number of columns and rows in the bundle, respectively). (d) Dense triangular packing of CNTs with minimum separation $\delta_{\text{min}}$ (Van der Waal’s gap) [54]. Varying density of metallic CNTs is modeled using $x > x_{\text{min}}$, where $x$ is the center-to-center distance between adjacent CNTs.

This paper provides a complete analysis of the applicability of SWCNT bundles at different domains (local, intermediate, and global) of VLSI interconnects. Some of the main contributions in this paper are CNT interconnect resistance modeling considering its dependence on all physical parameters, detailed discussion of factors affecting contact resistance of SWCNT interconnects, full 3-D capacitance simulation of CNT bundles for realistic VLSI interconnect dimensions, a quantitative discussion of the importance of inductive effects, analysis of the repeater power dissipation in buffered global interconnects, and an evaluation of CNT vias, which are already being fabricated by several research groups.

B. Equivalent Circuit Parameters for SWCNT Bundle Interconnects

The equivalent circuit model parameters for an isolated SWCNT of length smaller than the mean free path of electrons [51] are shown schematically in Fig. 3(a). This model is based on the Luttinger liquid theory [52], which describes interacting electrons in a 1-D conductor (since the commonly used Fermi liquid theory breaks down in 1-D systems [53]). Besides, the well-known magnetic inductance and electrostatic capacitance, the quantum resistance, quantum capacitance, and kinetic inductance are important considerations for low-dimensional systems such as CNTs. 

As mentioned in Section II, the high resistance of an isolated CNT necessitates the use of a bundle of CNTs [Fig. 3(c)] to form a low-resistance interconnect. The minimum intertube spacing observed in SWCNT bundles is 0.32 nm, characteristic of Van der Waals intertube binding [54]. This close-packed structure, shown schematically in Fig. 3(d), is the configuration considered in the analysis of CNT interconnects in this paper. In practical reality, not all CNTs of such a bundle are metallic. Moreover, other fabrication techniques do not yield such densely packed CNT bundles. Sparsely populated bundles and the presence of nonmetallic CNTs (which do not contribute to current conduction) are accounted for by taking a larger center-to-center distance between adjacent mCNTs ($x > x_{\text{min}}$ [Fig. 3(d)].

The electrical properties and equivalent circuit parameters for such a bundle need careful analysis in order to ascertain the applicability of CNTs as VLSI interconnects of the future. This
requires detailed modeling of the bundle resistance, capacitance, and inductance while incorporating fundamental physics of low-dimensional conductors. The models and their components are explained in detail in the following subsections.

1) SWCNT Interconnect Resistance Modeling:

a) Resistance of an isolated CNT: From the Landauer approach, the conductance of an isolated CNT can be modeled by \( G = 2(e^2/h)MT \) [24], where \( e \) is electronic charge, \( h \) is Planck’s constant, \( M \) is the number of allowed channels for transport, and \( T \) is the transmission probability. The factor of 2 accounts for spin degeneracy (spin up and spin down). An SWCNT has two conducting channels (\( M = 2 \)). Hence, the lowest possible resistance of an isolated SWCNT, assuming perfect metal–CNT contacts (\( T = 1 \)), is given by [24]

\[
R_Q = \frac{h}{4e^2} = 6.45 \text{ k}\Omega.
\] (1)

This is the fundamental quantum of resistance associated with an SWCNT of length less than the mean free path of electrons [55] (\( l < \lambda_{\text{CNT}} \)), which is typically 1 \( \mu m \) [8], [56], [57]. For such lengths, electron transport within the nanotube is essentially ballistic and resistance is independent of length. \( R_Q \) for a CNT is equally divided between the two metal–nanotube contacts (assuming identical contacts) on either side. However, for lengths \( l > \lambda_{\text{CNT}} \), scattering leads to an additional resistance of \( (h/4e^2)/\lambda_{\text{CNT}} \) [58], which yields a distributed resistance per unit length (also confirmed by experimental observations [57], [59])

\[
R_s(\text{p.u.l.}) = \frac{h}{4e^2} \left( \frac{1}{\lambda_{\text{CNT}}} \right).
\] (2)

i) Voltage bias dependence of SWCNT resistance: The resistance of a nanotube is known to be a function of the voltage bias across it. At high voltages, resistance increases due to enhanced phonon scattering of high-energy electrons [57]. The inset in Fig. 4 shows the monotonic increase in CNT resistance (\( R_{\text{CNT}} \)) as the voltage bias (\( V_{\text{bias}} \)) increases, following the relation [60]

\[
R_{\text{CNT}} = R_{\text{low-bias}} + \frac{V_{\text{bias}}}{I_0}
\] (3)

where \( R_{\text{low-bias}} \) is the CNT resistance at low bias (around 0 V) and \( I_0 \) is the saturation current (current does not increase beyond \( I_0 \) even if \( V_{\text{bias}} \) is increased). For a 1-\( \mu m \)-long CNT, \( I_0 \) is found to be around 25 \( \mu A \) [57], [60]. This is also the theoretically calculated value of \( I_0 \) [60]. However, the data in [57], [61] show that \( I_0 \) varies with CNT length.

In interconnect applications, the bias dependence of \( R_{\text{CNT}} \) can lead to an (instantaneous) increase in resistance. In order to quantify this effect, we first calculate \( V_{\text{Limit10\%}} \)—the largest value of \( V_{\text{bias}} \) for which the increase in \( R_{\text{CNT}} \) is less than 10\% of \( R_{\text{low-bias}} \)—using

\[
V_{\text{Limit10\%}} = I_0 (1.10 R_{\text{low-bias}} - R_{\text{low-bias}})
= 0.11I_0 R_{\text{low-bias}}.
\] (4)

b) Circuit schematic for SPICE simulation: Fig. 5 shows the maximum instantaneous voltage drop \( V_{\text{bias}} \) (solid vertical lines and square symbols) across a switching CNT bundle interconnect for different lengths (\( l \)) and driver sizes (\( s \)), from SPICE simulation at 22 nm node (\( V_{\text{DD}} = 1V \)). Numbers adjacent to squares are the number of metallic CNTs in the bundle. Curves show \( V_{\text{Limit}} \) for 10\% (solid), 20\% (dashed), and 30\% (dotted) increase in \( R_{\text{CNT}} \). Circles correspond to data points on Curve B in Fig. 4. Curves are produced using linear extrapolation since data beyond few \( \mu m \) lengths are not available. (b) Circuit schematic for SPICE simulation.

Based on published experimental data from various sources, Fig. 4 plots the calculated values of \( V_{\text{Limit10\%}} \). The use of experimental data allow us to account for the length dependence of \( I_0 \). It can be observed that the value of \( V_{\text{Limit10\%}} \) increases as CNT length increases. This is because longer CNTs have higher \( R_{\text{low-bias}} \).

Fig. 5 compares the actual instantaneous voltage drop across a CNT bundle interconnect, \( V_{\text{bias}} \), to the limiting values of voltage bias calculated for 10\%–30\% increase in CNT resistance. \( V_{\text{Limit10\%}} \) and \( V_{\text{Limit30\%}} \), \( V_{\text{bias}} \), for CNT interconnects of different lengths and driver sizes (for local, intermediate, and global levels), with varying number of CNTs per bundle, are
obtained from SPICE simulations for the configurations shown in Fig. 5(b). \( V_n \) are obtained from (4), corresponding to the data for curve B in Fig. 4. As an example, if \( V_n \) lies in the shaded region above the curve \( V_{\text{limit}10\%} \), the instantaneous increase in its resistance will be greater than 30%.

It is found that for very short interconnects (\( l = 300 \text{ nm} \), driven by minimum-sized drivers), as few as 50 CNTs in a bundle lead to a voltage drop less than \( V_{\text{limit}10\%} \). In other words, these interconnects will have less than 10% instantaneous increase in resistance. This is because most of the voltage drop occurs across the large driver resistance. For longer interconnects, with driver size five or ten times the minimum size, a few 100 CNTs in a bundle are needed to ensure that the instantaneous increase in resistance is less than 10%. For wide global interconnects, which can easily accommodate several thousands of CNTs, the worst-case effect of \( V_n \) on resistance is negligible.

Hence, for all practical purposes, the resistance of CNT bundle interconnects is independent of voltage. However, this may not be true for CNT monolayer interconnects \([42, 43, 45]\), which can accommodate fewer than 20 CNTs. It is important to note that the voltage drop across an interconnect is transient, and the instantaneous increase in resistance discussed here does not translate into a commensurate increase in interconnect delay.

ii) Imperfect metal–nanotube contact resistance (\( R_C \)): The observed dc resistance of a CNT (at low bias) is much higher than the resistance \( R_Q \) or \( R_S \) shown in the previous sections, due to the presence of imperfect metal–nanotube contacts. Experimental reports on the nature of contacts to CNTs have shown varied results, with Schottky contacts reported for Ti \([62]\) and ohmic contacts reported for Au \([63]\) and Pd \([61, 64]\). Experimental evidence shows that the nature of contacts also depends on nanotube diameter \([65]\). A rational function fit for the data in \([65]\), with nanotube diameter as a parameter, is shown in \([46]\). However, the metal–CNT contact properties depend on several other physical parameters, such as contact geometry (planar versus embedded contacts shown in Fig. 6), contact metal work function, and temperature.

A theoretical analysis of contacts between metals and semiconducting nanotubes is presented in \([66]\), showing the effect of weak band realignment due to charge transfer at the contacts. A transition in the nature of the contact from a Schottky barrier to an ohmic contact is shown as the diameter increases. The diameter at which this transition occurs, or the crossover diameter \( (d_{S-O}) \), is given by \([66]\) the following equation:

\[
d_{S-O} = d_0 \left( 1 + \alpha \sqrt{\frac{k_B T}{\phi_m - \phi_{\text{CNT}}}} \right) = d_0 \left( 1 + \frac{\delta d}{d_0} \right) \tag{5}
\]

where \( d_0 \) is the crossover diameter if no charge transfer were considered, \( k_B \) is Boltzmann’s constant, \( T \) is temperature, and \( \phi_m - \phi_{\text{CNT}} \) is the difference in work function between metal contact and the CNT. \( \alpha \) is a parameter that depends on the contact geometry (\( \alpha \approx 14 \) for planar contacts, \( \alpha \approx 2.5 \) for embedded contacts \([67]\)).

In order to provide good ohmic contacts to a wide range of nanotube diameters, \( d_{S-O} \) (and hence \( \delta d \)) must be small. Fig. 6 shows the effect of contact geometry, temperature, and metal work functions on the ratio \( \delta d/d_0 \). Evidently, it is desirable to have embedded contacts with a large difference in metal–CNT work functions.

For small-diameter metallic SWCNTs, the metal–nanotube chemical bonding significantly perturbs the CNT due to high curvature and chemical reactivity of these tubes, leading to a large tunneling barrier at the contacts \([66]\), which makes the contact resistance of metallic SWCNTs also diameter dependent. Both experimental \([65]\) and theoretical \([66]\) analyses have shown that it is very difficult to obtain good ohmic contacts for metallic SWCNTs with small diameter around 1 nm. Hence, even though small contact resistances have been achieved in some laboratory experiments \([56, 68]\), it is important to consider the possibility of large (imperfect) metal–nanotube contact resistance in the practical application of CNT interconnects.

The contact resistance \( R_C \) is divided equally between the two contacts and is in series with the quantum resistance \( R_Q \). The total resistance of an isolated SWCNT is given by the following:

\[
R_{\text{CNT}} = \begin{cases} R_C + R_Q, & \text{if } l << \lambda_{\text{CNT}} \\ R_C + R_Q + lR_S, & \text{if } l > \lambda_{\text{CNT}} \end{cases} \tag{6}
\]

Evidently, the resistance associated with a single CNT is too high for realizing an interconnection. Hence, it is necessary to have a bundle of CNTs to lower the effective resistance.

b) Resistance of an SWCNT Bundle Interconnect: The effective resistance of a CNT-bundle with \( n_{\text{CNT}} \) CNTs is given by

\[
R_{\text{bundle}} = \frac{R_{\text{CNT}}}{n_{\text{CNT}}}. \tag{7}
\]

It is assumed here that the interaction between adjacent CNTs of a bundle is weak. Experiments on CNT bundles have shown a large temperature-independent coupling resistance (\( \sim \) several
MΩ [59] for defect-free nanotubes, characteristic of direct tunneling between the nanotubes. Hence, the CNTs of a bundle are only weakly coupled and can be safely assumed to carry currents independent of each other.

i) Temperature dependence of CNT bundle resistance: The resistance of a 3-μm-long isolated metallic SWCNT has been shown to increase superlinearly with temperature from about 50 to 290 K [69]. This nonlinearity has been attributed to optical phonon absorption in a temperature-dependent model in [70] with very large increase in resistance predicted for longer CNTs in the range of 400 K (typical on-chip temperatures), although this is not supported by data. On the other hand, measurements on SWCNT bundles have consistently shown linear increase in resistance with temperature [71], [72] for temperatures as high as 580 K over a wide range of CNT bundle lengths.

The calculated rate of change of resistivity with temperature for a densely packed SWCNT bundle [72] is \( \frac{\rho}{\rho_0} = 0.005 \mu\Omega \cdot \text{cm/K} \), while experimental measurements have shown \( \frac{\rho}{\rho_0} \sim 0.1 \mu\Omega \cdot \text{cm/K} \) [72]. The discrepancy apparently arises from the presence of semiconducting CNTs as well as from inter-CNT interactions. Fig. 7 shows that, for these values of \( \frac{\rho}{\rho_0} \), SWCNT bundle resistance increases by about 10% from room temperature to 400 K (Cu resistance increases by about 40% in the same temperature range).

ii) Diameter dependence of CNT bundle resistance: Semiconducting CNTs have a large chirality-induced primary bandgap, which is inversely proportional to nanotube diameter \( d \), as well as a curvature-induced secondary bandgap, which scales as \( 1/d^2 \) [73]. Semimetallic nanotubes have a very small curvature induced bandgap, which scales as \( 1/d^2 \). Armchair nanotubes (chiral indexes \( m = n \)), however, are truly metallic as they have zero bandgap, independent of diameter [74].

The small diameter of SWCNTs, and the absence of a controlling gate bias in interconnect applications, means that only the metallic nanotubes in an SWCNT bundle interconnect contribute to conduction. Hence, the variation in bandgap with nanotube diameter will not affect the performance of SWCNT bundle interconnects.

The electron mean free path of a CNT, which has a large impact on CNT resistance, increases linearly with nanotube diameter [75]. In the scope of this study, we assume that all single-walled nanotubes have 1 nm diameter and a mean free path, \( \lambda_{\text{CNT}} = 1 \mu m \), the typical value reported in the literature.

Fig. 8(a) shows that at the local level, for mCNT density as low as 30%, CNT bundle resistance can be lower than that of Cu interconnects when \( R_C \) is low. With high \( R_C \), higher mCNT densities are needed. Moreover, as long as the CNT bundle resistance is less than that of Cu, the resistance can also be considered independent of the bias voltage (falling outside the shaded regions). With technology scaling, fewer nanotubes can be accommodated in the smaller cross sections, thereby increasing the bundle resistance, but the Cu interconnect resistance also increases proportionately. Hence, the same trends in comparison with Cu are seen for 22 nm node and 14 nm node (inset).

However, as the shaded region of the plot expands at advanced technology nodes, higher mCNT densities will be necessary as technology scales in order to keep bundle resistance independent of voltage bias conditions.

Fig. 8(b) shows that for densely packed bundles at the global interconnect level, the bundle resistance is less than that of Cu for a wide range of interconnect lengths, even for low mCNT density (45%). \( R_C \) as large as 10 kΩ does not have a large impact on the overall bundle resistance, while the relative impact of \( R_C = 100 \text{ kΩ} \) also becomes small for long interconnects. This is because \( R_C \) is independent of length whereas the scattering resistance component \((lR_S)\) increases linearly with length, as shown in (6). The resistance of maximum width global interconnects (width \( \sim 1 \mu m \)) with dense CNT bundles stays in the range of 40%–50% of Cu global interconnects and does not change appreciably over length or with technology [Fig. 8(c)]. Even with imperfect metal–CNT contacts, CNT bundle resistance \( \sim 50\% \) of Cu wire resistance is easily achieved for long, wide global interconnects.

2) SWCNT Interconnect Capacitance Modeling: An external voltage applied to a 1-D conductor affects its electrochemical potential energy in two ways. The change in electrostatic potential energy when a charge \( \delta Q \) is added to the conductor is given by \( \delta F = \frac{\delta Q^2}{2C_G} \), where \( C_G \) is its electrostatic capacitance. Due to the low density of states at the Fermi energy level, the charge \( \delta Q \) has to occupy available quantum energy states above the Fermi energy level. The additional quantum energy required to add charge \( \delta Q \) to these higher energy states can be modeled as \( \delta F = \frac{\delta Q^2}{2C_G} + \frac{\delta (Q^2)}{2C_Q} \), where \( C_Q \) is its quantum capacitance [17]. Hence, the net change in electrochemical potential energy is given by \( \delta F = \frac{\delta Q^2}{2C_G} + \frac{\delta (Q^2)}{2C_Q} \), which indicates that the effective capacitance of a 1-D conductor is given by a series combination of \( C_E \) and \( C_Q \) [Fig. 9(a)].

The quantum capacitance depends on the density of states at the Fermi energy level \( D(\mu) \). For a 1-D conductor, \( D(\mu) = 1/(\pi \hbar v_F) \), where only one spin direction for electrons has been taken into account. Here, \( v_F \) is the Fermi velocity \( v_F \approx 8 \times 10^5 \text{ m/s} \) and \( 2\pi \hbar = h \), the Planck’s constant. Hence, \( C_Q \) is given by [17] and [51]:

\[
C_Q(\text{p.u.l.}) = e^2 D(\mu) = \frac{e^2}{\pi \hbar v_F} = \frac{2e^2}{\hbar v_F}.
\]
Fig. 8. (a) CNT bundle local interconnect resistance (length = 1 µm) as a function of metallic CNT (mCNT) density (normalized to maximum density), at 22 nm node (main figure) and 14 nm node (inset). Length = 1 µm, hence CNTs are assumed to be ballistic. Bundle dimensions and Cu wire resistance for identical dimensions are shown. Temperature = 300 K. In shaded regions, resistance may become dependent on voltage bias (see Fig. 5). Number of CNTs in bundle is shown with arrows. (b) Resistance of dense CNT bundle minimum-width global interconnect as a function of length at 22 nm node, for different values of \( R_C \) and m-CNT density. (c) Ratio of densely packed CNT bundle resistance to Cu wire resistance for maximum-width global wire (\( w \sim 1 \mu m \)), as a function of interconnect length and technology node. Temperature = 300 K for all figures.

Note that for conventional 3-D conductors, the density of states \( D(\mu) \) is large; hence, \( C_Q \) is large and its effect on overall capacitance is negligible.

Since an SWCNT has two conducting subbands at the Fermi energy with two electronic spin directions, each SWCNT has four 1-D conducting channels [47]. The net quantum capacitance of a CNT is given by a parallel combination of four \( C_Q \) elements, thus

\[
C_Q^{\text{CNT}}(p.u.l.) = 4C_Q(p.u.l.) \approx 400 \text{ aF/µm}. \tag{9}
\]

In order to calculate the equivalent capacitance of an SWCNT bundle, we assume that the \( n_{\text{CNT}} \) nanotubes in the bundle are noninteracting and carry equal currents independent of each other [59]. Under this assumption, the SWCNT bundle is composed of \( 4n_{\text{CNT}} \) 1-D conductors (each nanotube contributes four 1-D channels). For an SWCNT bundle, the effective quantum capacitance is given by

\[
C_Q^{\text{bundle}} = 4n_{\text{CNT}}C_Q. \tag{10}
\]

The electrostatic capacitance of a CNT, \( C_E^{\text{CNT}} \), can be found by considering it as a thin cylindrical wire and is dependent on its surrounding environment. For the configuration shown in Fig. 10, the expression for \( C_E \) (per unit length) is given by [51]

\[
C_E^{\text{CNT}}(p.u.l.) = \frac{2\pi \varepsilon}{\cos \hbar \cdot (y/d)}. \tag{11}
\]

For \( d = 1 \text{ nm}, \ y = 1 \mu \text{m}, \ C_E^{\text{CNT}} \approx 30 \text{ aF/µm}.\)

a) Electrostatic Capacitance of an SWCNT Bundle: Fig. 11(a) shows a cross section of the interconnect structure considered for electrostatic analysis of mCNT bundle interconnects. Two immediately adjacent wires (left and right, held at ground potential) parallel to the CNT bundle are considered. Top and bottom ground planes represent the orthogonal metal layers, although electrostatic coupling to adjacent (left/right) interconnects is dominant for wires with aspect ratio \( t/w > 1. \) For
the electrostatic analysis, each CNT is treated as a metal with equal potential over the tube, similar to the approach in [76]. For the sake of simplicity, it is assumed that all CNTs are held at the same potential. Contrary to the claim in [43], this assumption is valid only when the contacts to all nanotubes within a bundle are identical and each CNT has the same mean free path. Under this assumption, the coupling capacitance between adjacent CNTs in the same bundle does not come into play.

Fig. 11(b) shows the total electrostatic capacitance with respect to left/right neighbors and top/bottom ground planes for each CNT in the bundle. These numbers are obtained from 3-D simulations on square-cross-section wires of dimensions equivalent to the 1 nm nanotubes using FastCap [77]. The numbers clearly show that inner nanotubes in the bundle are electrostatically shielded from the adjacent ground conductors. The edge nanotubes are the dominant contributors to the effective electrostatic capacitance of the CNT bundle. Hence, in order to facilitate extensive simulation of large CNT bundles, the shielded inner nanotubes can be ignored.

It must be noted that semiconducting nanotubes will always be present in a CNT bundle. Since no gate-bias voltage is applied to an interconnect, we can assume that the semiconducting CNTs are insulating. Hence, capacitance between metallic and semiconducting CNTs within a bundle will not be important. In this scenario, the semiconducting CNTs can still act as additional dielectric boundaries modifying the inter-mCNT capacitance of the bundle. However, this capacitance between mCNTs of the same bundle is also not important since they are assumed to be equipotential. The amount of charge injected into the semiconducting CNTs at low bias and the equipotential assumption for mCNTs in the bundle are all dependent on the quality, material, and uniformity of the metal–nanotube contacts. Although these may be important considerations for high accuracy, we ignore these modifications for the sake of simplicity.

Fig. 12 compares the net electrostatic capacitance of densely packed CNT bundles with that of equivalent-dimension Cu wires with respect to neighboring ground conductors. The dimensions of interconnects are kept small due to simulator time and memory limitations (other geometrical parameters shown in Fig. 11(a) are proportionately scaled). The results show that the capacitance of dense CNT bundles is lower (but within 10%) than that of Cu wires, for different values of aspect ratio.

Fig. 13 shows the comparison between CNT bundles and Cu wires of realistic dimensions below the 22 nm technology node. In the case of maximally dense bundles, the CNT bundle electrostatic capacitance is still found to be within 10% lower than that of Cu. However, as the mCNT density is reduced (assuming uniform distribution of mCNTs in the bundle), the capacitance becomes significantly lower than that of Cu interconnects.

It is important to note that the use of accurate field solvers for calculating capacitance is limited to configurations with relatively small number of conductors due to memory and time limitations (e.g., in Fig. 13, we could not obtain results for a completely dense bundle at 22 and 18 nm nodes due to the large number of CNTs in the bundle). This limitation becomes worse for wide global wires, which can accommodate thousands of CNTs. Hence, in the delay analysis of wide global interconnects in Section IV, we will use the earlier deduction that Cu wire capacitance provides an upper-bound (within 10%) for the capacitance of a dense CNT bundle. This deduction is supported by the results in Fig. 12 as well.

The effective capacitance of an SWCNT bundle interconnect ($C_{\text{bundle}}$) is given by the series combination of its electrostatic capacitance and quantum capacitance. As shown in Fig. 14, the effective SWCNT bundle capacitance is nearly equal to its electrostatic capacitance and the effect of the quantum capacitance is small.

3) SWCNT Interconnect Inductance Modeling:

a) Inductance of an isolated SWCNT: The inductance of a conductor models the energy associated with the motion of electrons carrying current ($I$) through it. The energy stored in the magnetic field generated by the current is given by $(1/2)L_M I^2$, where $L_M$ is its magnetic inductance. For a net current to flow through a conductor, there has to be an excess of electrons moving in the direction opposite to current flow. In a 1-D conductor that has low density of states at the Fermi energy level ($E_F$), these excess electrons can only be added at available quantum energy states above $E_F$, and thus have a higher kinetic energy. This additional kinetic energy, stored in the moving electrons responsible for current flow in a 1-D conductor, is modeled...
The magnetic inductance associated with an isolated SWCNT can be calculated from the magnetic field of a current-carrying wire, diameter “d,” lying a distance “y” away from a ground plane, as depicted in Fig. 10. Thus,

$$L_M^{\text{CNT}}(\text{p.u.l.}) = \frac{\mu}{2\pi} \cos^{-1} \left( \frac{y}{d} \right).$$  \hspace{1cm} (14)$$

For $d = 1$ nm and $y = 1$ µm, $L_M^{\text{CNT}}$ (per unit length) $\approx 1.5$ pH/µm. Clearly, kinetic inductance dominates the magnetic inductance for an SWCNT.

The equivalent inductance of a CNT is given by a series combination of its magnetic and kinetic inductances, as shown in Fig. 15(b). In familiar 3-D conductors with number of conducting modes $n \gg 1$, the kinetic inductance $L_K/n$ becomes small compared to the magnetic inductance, hence its effect becomes less evident [47].

While considering the importance of CNT inductance, it must be noted that the high intrinsic resistance of a CNT (6.45 kΩ/µm, considering perfect contacts) makes the inductive impedance appear small in comparison [Fig. 15(c)].


2) Inductance of an SWCNT Bundle Interconnect: In this paper, we have assumed noninteracting nanotubes since very little is known about the nature of the electromagnetic interactions in dense CNT bundles. Although SWCNTs in a bundle may not be isolated from each other, the consequences of this for the density of states, conductivity, and magnetically induced currents are not known yet. While the high-frequency properties of individual nanotubes as well as capacitive interactions between adjacent nanotubes in a flat array have been studied [76], [80], to the best of our knowledge, there is no experimental work or theoretical analysis yet about the nature of electromagnetic interactions between nanotubes or the penetration of electromagnetic waves inside a nanotube. Hence, the mutual inductance between SWCNTs in a bundle merits rigorous investigation before these effects can be included in the inductance model. The only work in the existing literature that attempts to calculate the mutual inductance between the nanotubes of a CNT bundle interconnect [48] assumes a nanotube as a solid metallic conductor with an equivalent conductivity. Since there is no physical justification for the validity of this approach, we refrain from using this model. In the following section, we show that CNT bundle inductance does not have a significant impact on the performance of CNT-based interconnects. Hence, neglecting the mutual inductance between CNTs in a bundle will not have a large impact on the performance analysis shown here.

Under the assumption that the nanotubes of a bundle are magnetically isolated (no mutual inductance effects), the effective inductance can be treated as the parallel combination of the inductance corresponding to each SWCNT. The inductance of the CNT bundle interconnect is then given by

$$L_{\text{bundle}}(\text{p.u.l.}) = \frac{L_K^{\text{CNT}}(\text{p.u.l.}) + L_M^{\text{CNT}}(\text{p.u.l.})}{n_{\text{CNT}}},$$

The scaling of bundle inductance with number of CNTs is also supported by experimental observations [79].
C. Importance of SWCNT Bundle Inductance for Interconnect Delay

The importance of considering CNT bundle inductance in interconnect delay models can be analyzed following the methodology in [81], which presents a detailed analysis of resistance–inductance–capacitance (R–L–C) representation of interconnects. An R–C-only model for interconnect delay becomes inaccurate (and inclusion of inductance in the delay model becomes necessary) when

\[ Z_{dr}C_l < \frac{1}{2} R_l C_l < Z_0 C_l. \]

(16)

Here, \( Z_{dr} \) is the driver impedance; \( R, C, \) and \( L \) are the per unit length interconnect resistance, capacitance, and inductance, respectively; \( l \) is the interconnect length; and \( Z_0 = \sqrt{\frac{L}{C}} \) is its characteristic impedance. The first inequality in (16) suggests that the gate (driver) delay should be less than the interconnect delay (approximated by \( \frac{1}{2} R_l C_l \)) as well as wave propagation delay (\( Z_0 C_l \)). The second inequality suggests low resistive losses (\( R_l < 2Z_0 \)) to ensure low-loss transmission line behavior. If these conditions are satisfied, the interconnect will behave as a transmission line, with delay determined by the wave propagation speed \( 1/\sqrt{LC} \), and its inductance will be important for accurate delay estimation.

Fig. 16 plots the three terms in (16) for different CNT bundle interconnect geometries. It is observed that for all different interconnect lengths and dimensions, as well as driver sizes (chosen based on intermediate or global interconnect levels), either the driver impedance (\( Z_{dr} \)) or the line impedance (\( R_l \)) dominates. This remains true for minimum-width wires as well as for the widest global interconnects. Since the inequality (16) is never satisfied, it is reasonable to assume that simple RC circuits (ignoring line inductance) will suffice to model CNT bundle interconnect delay.

It must be noted that the mutual inductance between nanotubes within a CNT bundle has not been taken into account here. However, we overestimate the magnetic inductance of each CNT (by considering the substrate as the nearest current-return path) as well as kinetic inductance. Hence, a more accurate calculation of CNT bundle inductance will not alter the conclusions shown here.

D. Equivalent Circuit Model for SWCNT Bundle Interconnect Delay

The equivalent circuit model for an SWCNT bundle interconnect is shown in Fig. 17. The driver is represented by its equivalent resistance \( R_{dr} \), input capacitance \( C_{\text{input}} \), and output parasitic capacitance \( C_p \), while \( C_{\text{LOAD}} \) represents the input capacitance of the driven gate. Two types of interconnects are analyzed here: 1) local interconnects with driver size one to five times the minimum-sized inverter and a load fan-out of four identical gates and 2) global interconnects where the interconnect is part of a long buffered interconnect, with identical (large) inverters at driver and load ends. As shown in the previous subsection, CNT bundle inductance has little impact on the interconnect delay, hence a distributed RC model is used to represent the interconnect.
As an illustration, the commonly used Elmore delay [82] expression for a Cu interconnect is given by

\[ t_{50\%}^{\text{Cu}} = 0.69R_{\text{dr}}(C_{\text{LOAD}} + C_p) + 0.69R_{\text{dr}}Cl + 0.38RC_l^2 + 0.69R\frac{C_{\text{LOAD}}}{C_p}. \]  

(17)

where \( R \) and \( C \) are the interconnect resistance and capacitance per unit length, respectively, and \( l \) is the interconnect length. For local interconnects, the drivers and load are nearly minimum-sized gates, hence driver resistance \( R_{\text{dr}} \) is large, while load capacitance \( C_{\text{LOAD}} \) is small. On the other hand, the buffers/repeaters needed to send signals across large distances are sized up significantly (often more than 100 times minimum-sized gates). Hence, for global interconnects, \( R_{\text{dr}} \) is small while \( C_{\text{LOAD}} \) is large. Similar Elmore delay expressions can be obtained for the equivalent circuit shown in Fig. 17(b) as follows:

\[ t_{50\%}^{\text{CNT}} = 0.69R_{\text{dr}}(C_{\text{LOAD}} + C_p) + 0.69R_{\text{dr}}Cl \]

\[ + 0.69\left(R_{\text{dr}} + \frac{R_C + R_Q}{2n_{\text{mCNT}}}\right)\left(\frac{C_{\text{bundle}}}{C_Q} + C_{\text{bundle}}\right)l \]

\[ + 0.38\frac{R_S}{n_{\text{mCNT}}}\left(\frac{C_{\text{bundle}}}{C_Q} + C_{\text{bundle}}\right)l^2 \]

\[ + 0.69\left(\frac{R_S}{n_{\text{mCNT}}} + \frac{R_C + R_Q}{n_{\text{mCNT}}}\right)C_{\text{LOAD}}. \]  

(18)

IV. INTERCONNECT PERFORMANCE ANALYSIS WITH SWCNT BUNDLES

A. Local Interconnects

Local interconnect lengths are typically a few hundred nanometers and no more than a few microns. Hence, for this analysis, it is assumed that local interconnect lengths are less than the mean free path, \( \lambda_{\text{CNT}} \). Fig. 18 shows the propagation delay of short local interconnects that are driven by minimum-sized drivers. As explained in Section III-B2, since full 3-D capacitance simulations are used to accurately compute the SWCNT bundle electrostatic capacitance, the delay at 22 and 18 nm nodes is calculated only for sparse bundles (mCNT density <100%) with fewer SWCNTs. It is found that the propagation delay of short local CNT interconnects is nearly the same as that of Cu wires (delay ratio \( t_{\text{CNT}}/t_{\text{Cu}} \approx 1 \)) over a wide range of mCNT densities and for all technology nodes considered. This is despite the fact that sparse bundles have higher resistance than Cu wires, since the driver resistance (\( \sim k\Omega \)) is much larger than the resistance of short local interconnects. In fact, as shown in Fig. 18(d), even with imperfect contacts (\( R_C = 10\ k\Omega \)) that increase the CNT bundle resistance further, delay remains competitive with that of Cu wires.

As noted before, due to lack of control on chirality, it is difficult to achieve high density of metallic SWCNTs in a bundle. It is shown here that even with the low densities that may be readily achievable (\(~ 40\% \) mCNTs are routinely fabricated), SWCNT bundles can compete with the performance of Cu for short local interconnects.

In the case of very short interconnects (<1 \( \mu m \) length), where \( R_{\text{dr}} \) is much greater than interconnect resistance, it is beneficial to make the SWCNT bundles sparse, which reduces interconnect capacitance (see Fig. 13). CNT monolayer interconnects proposed in [45], which consist of a single flat array of SWCNTs [Fig. 3(b)], are an extreme example of this effect. The resistance of a monolayer is high because it can accommodate only a small number of SWCNTs, but its capacitance is nearly half that of a Cu wire of equal width. It has been previously reported that CNT monolayers are suitable for local interconnects up to 20 \( \mu m \) lengths [43]. We show here that the realistic range of lengths where CNT monolayers may be useful is significantly smaller and depends on the driver size.

1) CNT Monolayer Interconnects: Table I shows the driver resistance for two different driver sizes and interconnect resistances considering Cu as well as SWCNT monolayer interconnects at the 22 nm node. In this case, \( R_C = 0 \), assuming high-quality embedded contacts as done in [43]. It is important to note that ensuring embedded contacts (shown schematically in Fig. 6) for each SWCNT makes it difficult to pack SWCNTs with maximum density since the embedded contact material will occupy some inter-CNT space. We have optimistically assumed that ten metallic SWCNTs (diameter 1 nm each) can be

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
\multicolumn{2}{|c|}{\text{Driversize}} & \multicolumn{2}{|c|}{\text{Minimumsize}} & \multicolumn{2}{|c|}{\text{Minimumsize}} \\
\hline
\text{CuResistance(Ohms)} & 124.2 & 496.8 & 745.2 & 993.6 & 1242 \\
\hline
\text{CNTmonolayerResistance(Ohms)} & 1290 & 5160 & 7740 & 10320 & 12900 \\
\hline
\end{tabular}
\caption{Driver Resistance for Different Driver Sizes and Interconnect Resistances for Different Length Interconnects with Cu and CNT Monolayers at 22 nm Node}
\end{table}
accommodated in a 22-nm-wide monolayer interconnect. It is evident from Table I that few-micrometer-long Cu interconnects have a small resistance compared to the driver. However, this is not true for CNT monolayer interconnects unless the driver is minimum-sized.

Fig. 19 shows a comparison between Cu and CNT monolayer local interconnect delays for two different driver sizes (with a fan-out of four identical gates). For minimum-sized drivers (hollow symbols), CNT monolayers have lower delay than Cu because of the significantly lower capacitance. However, minimum-sized drivers can be used only when interconnect delay is dominated by the intrinsic gate delay. It is observed that at about 1 \( \mu \)m length, the local interconnect delay becomes much larger than intrinsic gate delay. Beyond this length, it is necessary to use larger drivers. There is a clear reduction in delay with driver size ten times the minimum (solid symbols). However, in this case, the large resistance of CNT monolayers increases delay by 100%–200% as compared to Cu, in spite of the 2× lower capacitance.

Hence, CNT monolayers can provide only some delay advantage over Cu interconnects for very short, a few hundred nanometers long, local interconnects that are driven by minimum-sized drivers. For this small subset of interconnects, delay is dominated by the intrinsic gate delay. The advantage for CNT monolayers is the reduced power dissipation resulting from 2× lower capacitance [42], [43], [45]. However, it must also be noted that the resistive power dissipation along these interconnects will be higher than that across Cu interconnects.

B. Global Interconnects

Fig. 20 shows that dense SWCNT bundle interconnects can easily outperform Cu at the global level (ratio of delay \( t_{\text{SWCNT}} / t_{\text{Cu}} < 1 \)). The improvement in global interconnect performance with CNT bundles arises mainly from the large mean free path of electrons, which results in much lower interconnect resistance [see Fig. 8(b) and (c)]. In the case of minimum-width global interconnects, dense CNT bundles can reduce the delay by more than 55% for a 500-\( \mu \)m-long interconnect. In the case of maximum-width global interconnects whose width is \( \sim 1 \mu \)m at all technology nodes, dense SWCNT bundle interconnects can reduce global interconnect delay by 30% for a 1-mm-long and 40% for a 3-mm-long interconnect. The larger improvement in case of minimum-width wires is because of higher Cu resistance for these dimensions. The benefit of using SWCNT global interconnects increases with scaling and with the interconnect length. As mentioned in Section III-B2, SWCNT bundle electrostatic capacitance at the global level is assumed to be equal to Cu wire capacitance as it provides a close upper bound to the CNT bundle capacitance.

Global interconnects are often designed by inserting buffers (repeaters) at regular intervals to minimize delay per unit length \( (\tau / l) \) [82]. Fig. 21 shows the optimal buffer insertion parameters: 1) \( l_{\text{opt}} \), the optimal distance between repeaters and 2) \( s_{\text{opt}} \), the optimal size of repeaters as a multiple of a minimum-sized inverter. Fig. 21(c) shows that the optimal delay per unit length \( (\tau / l)_{\text{opt}} \), with buffered SWCNT bundles as minimum-width global interconnects, is significantly lower than that with Cu. While technology scaling alone shows a small increase in minimum-width Cu global interconnect delay, the move from Cu to CNT-based interconnect reduces delay by nearly 2×.

Fig. 22 shows the power dissipation of a 1-mm-long buffered global interconnect as a multiple of the power dissipation of a minimum-sized inverter (the ratio of the total width of buffers driving the interconnect to the width of a minimum-sized inverter, which is shown in the figure, is commensurate to the ratio of their power dissipations). It is found that the power dissipation of a minimum-width optimally buffered global interconnect is the same with SWCNT bundle as that with Cu at all technology nodes [Fig. 22(a)]. Note that in this optimal-delay case, SWCNT interconnects have nearly 2× lower delay.
than Cu. Fig. 22(b) shows the power dissipation when the delay with SWCNT bundles is suboptimal but equal to the optimal delay with buffered Cu interconnects. In this case, the power dissipation with SWCNT bundle is $4 \times$ lower at the 22 nm node and nearly $8 \times$ lower at the 14 nm node. This result will have significant implications for future technology generations where power dissipation (active as well as leakage) is a critical concern.

In order to reduce delay, timing-critical global wires are conventionally kept wide ($\sim 1 \, \mu m$) and not scaled aggressively [14]. Fig. 23(a) shows that for such wide wire lengths, the optimal delay is much lower than that for minimum-width wires [shown in Fig. 21(c)], while the delay with SWCNT bundles is again significantly lower than that with Cu. However, Fig. 23(b) shows that the reduction in delay by using wide wires comes with an increase in power dissipation [in comparison with Fig. 22(a)]. Fig. 23(b) also shows that for 1-$\mu m$-wide global interconnects, SWCNT bundles can achieve the same delay as optimally buffered Cu interconnect with significantly lower power dissipation. Hence, SWCNT bundles can be used to improve the performance as well as to reduce the power consumption of global level VLSI interconnects.

V. INTERCONNECT RELIABILITY ANALYSIS WITH SWCNT BUNDLES

Due to strong $sp^2$ bonding (like graphite), CNTs are much less susceptible to electromigration (EM) problems than Cu interconnects. Metallic SWCNT bundles can carry current densities of the order of $10^9 \, A/cm^2$ [23]. The maximum current density in Cu interconnects, on the other hand, is limited to $\sim 10^6 \, A/cm^2$ due to EM [83]. A Cu interconnect with 100 nm × 50 nm cross section can carry current up to 50 $\mu$A, whereas a 1-nm-diameter CNT can carry up to 20–25 $\mu$A [57]. Hence, from a reliability perspective, a few CNTs are enough to match the current-carrying capacity of a typical Cu interconnect. It is only the need to reduce CNT interconnect resistance (and make its delay competitive with Cu) that necessitates bundles of a large number of CNTs. Fabrication efforts for CNT-based interconnects have thus far focused on vias, or vertical interconnects, which are most prone to failure because of their small dimensions [8], [19]–[22]. Here, we point to another property of CNTs that makes them attractive candidates for future vias.

At nanometer dimensions, interconnect metal layers in the back end see very significant temperature rise above the chip’s junction temperature as a combined result of increasing Cu interconnect resistivity and higher current density ($J$) demands [15]. This adversely affects interconnect reliability, since EM lifetime of Cu depends quadratically on $J$ and exponentially on metal temperature ($T_m$) [83]. The primary mode of heat removal from the interconnect stack (10–14 metal layers) is by conduction toward the heat sink attached to the chip substrate [15]. Due to low thermal conductivity ($K_{th}$) of interlayer dielectric (ILD) ($K_{th, ILD} < 0.4 \, W/mK$, and decreases further with technology), the Cu vias (with $K_{th, Cu} = 385 \, W/mK$) between successive metal layers form the primary heat conduits.

Molecular dynamics simulations of CNTs predict that $K_{th} \sim 6600 \, W/mK$ [5], an order of magnitude higher than that of Cu. The thermal conductivity of individual MWCNTs, found experimentally, exceeds 3000 W/mK [84]. The thermal conductivity of CNT bundles, however, can be much lower due to the presence of numerous highly resistive thermal junctions between adjacent nanotubes [84]. Based on measurements, $K_{th}$ for an SWCNT bundle is predicted in the range 1750–5800 W/mK [85] (along the CNT length) at room temperature. Hence, vias composed of CNT bundles can serve as much more effective heat conduits than Cu vias and reduce the temperature gradient in the back-end.

For SWCNT bundles, the high value of $K_{th}^{C_{th}}$ is only in the longitudinal direction (see Fig. 24). In the transverse direction, $K_{th}^{C_{th}}$ is about 20 times lower [86]. The $K_{th}$ of 1-$\mu m$-long CNT bundles was measured in [85]. Few experimental data and thermal models are available for $l < 1 \, \mu m$, which will be the desired
height of vias in most nanoelectronic circuits. The work in [70] suggests temperature and length-dependent models for $K_{th,h}$, of isolated metallic SWCNTs, but the same models may not be applicable to CNT bundles.

The results shown here are based on the assumption that the high thermal conductivities mentioned above can be achieved with CNT bundle vias, although the presence of large thermal contact resistance between CNT bundles and surrounding materials make it challenging. It is also assumed that the CNT bundles have high density and low electrical contact resistance, so as to have resistance similar to Cu vias (this requirement is further examined in the following subsection). For low via resistance values, self-heating in individual vias has a small impact on the overall temperature of the interconnect stack [15]. It has been shown in [32] that CNT vias are both electrically and thermally ballistic and heating is primarily due to the contacts. Note that we do not consider severe Joule heating effects at high bias [87] as the voltage bias across interconnects is very low. The impact of thermal contact resistance to SWCNTs [32], [88] has not been considered in this analysis. Hence, the vias are treated essentially as heat conduits raising the effective thermal conductivity between successive metal layers.

A. Hybrid CNT–Cu Interconnects

Full 3-D electrothermal simulations are performed on a complete 10–14 metal layer interconnect stack at different technology nodes, considering CNT-based vias forming vertical interconnections between Cu metal layers. This configuration is termed hybrid CNT–Cu interconnects. Fig. 25 shows that with these hybrid interconnects, the maximum temperature rise on the interconnect stack can be almost ten times smaller than that with traditional Cu interconnects.

Fig. 26(a) compares the resistance of local vias composed of CNT bundles to those made of Cu as a function of mCNT density at the 22 nm node. It is found that the resistance of CNT bundle vias can be five to ten times that of Cu vias even for very-high-density metallic-CNT bundles. Large imperfect metal–nanotube contact resistance ($R_C$) will have an adverse effect on the resistance of these vias. Hence, it is very important to achieve near-perfect contacts to keep CNT bundle via resistance low. Fig. 26(b) shows the delay penalty incurred by using vias with high resistance. For very short local interconnects with minimum-sized drivers, the delay penalty is less than 1% if the via resistance is less than 20 times Cu via resistance. As indicated in the figure, resistance in this range has already been achieved for MWCNT vias [8], [20], [21]. However, at the global level (where driver resistances are significantly smaller), resistance must be within four times that of Cu vias in order to limit delay penalty within 1%. This analysis assumes that SWCNT-based vias are used between all metal layers, and does not use multiple parallel vias to reduce the resistance.

B. Impact on Cu Interconnect Performance and Reliability

Fig. 27 shows the performance and reliability impact of using hybrid CNT–Cu interconnects. Since Cu resistivity increases with temperature, delay per unit length ($\tau/l$) for Cu global interconnects at maximum temperature is about 40% higher than optimal $\tau/l$ calculated at room temperature if Cu vias are used. This number is only about 10% higher if CNT bundle vias are used [Fig. 27(a)]. Fig. 27(b) shows that the lower interconnect temperatures with CNT bundle vias can lead to two orders of magnitude improvement (for $K_{th} = 5800$ W/mK) in the EM
mean-time-to-failure of Cu global interconnects and one order of magnitude improvement with $K_{th} = 1750$ W/mK.

Fig. 28 summarizes a thermal management strategy based on the use of CNT bundles to replace Cu at different locations in the interconnect stack. Full 3-D finite-element electrothermal simulation results are shown for the back-end thermal profile employing hybrid CNT–Cu interconnects. The variation in this profile and the maximum interconnect temperature is shown as the location of CNT bundles is changed from local to intermediate to global level vias. It is found that CNT bundle vias at the intermediate level yield the largest benefit in reducing the interconnect temperatures with least disruption in the Cu interconnect stack. This is because the intermediate level has the maximum number of metal layers, leading to the largest reduction in overall thermal resistance between the global metal layers and the heat sink. Thus, vias composed of CNT bundles can very effectively be used as thermal vias to reduce interconnect temperatures at advanced technology nodes, as well as in 3-D ICs where thermal management is a critical concern [89], [90].

VI. CONCLUSION

In conclusion, a comprehensive and realistic evaluation of SWCNT bundle interconnects has shown significant advantages over Cu in terms of performance, power dissipation, as well as thermal management/reliability. The equivalent circuit parameters needed for analyzing SWCNT interconnect delay have been studied in detail, highlighting the underlying assumptions. The correct length-dependent behavior of the voltage bias-dependent SWCNT interconnect resistance and the factors influencing CNT interconnect contact resistance have been presented. The first full 3-D capacitance simulations for SWCNT bundles of realistic VLSI interconnect dimensions, and a quantitative analysis of the importance of CNT interconnect inductance in computing interconnect delay, are shown. In light of the significant progress in state-of-the-art on-chip integration of CNT vias (vertical interconnects), this study provides a unique insight into the advantages and limitations of SWCNT bundle vias. The most important conclusions from this study are summarized below.

It is found that SWCNT bundles can provide a $4 \times$ reduction (at 22 nm node) in power dissipation of long global interconnects while offering the same performance as optimally buffered Cu interconnects. This power saving increases with technology scaling to $8 \times$ at the 14 nm node. This is a promising result as the rising power dissipation in aggressively buffered Cu global interconnects threatens to become a tremendous challenge for future technology generations. On the other hand, for performance-critical applications, SWCNT interconnects offer 30%–40% improvement in the delay of millimeter-long global interconnects as compared to Cu. The factor limiting SWCNT bundle performance in this domain is the density of mCNTs, not imperfect contact resistances. In the case of short local interconnects that have lengths less than the mean free path of electrons in a nanotube, SWCNT bundle performance again exceeds that of Cu. In this domain, the limiting factors are mCNT density as well as imperfect metal–nanotube contact resistance. In the case of vias, reliability is a far greater concern than performance. Hence, even with little or no improvement in signal propagation delay, SWCNT bundles can provide a valuable alternative to Cu due to better reliability. Most importantly, it is shown that SWCNT bundle vias can greatly reduce interconnect temperature rise and, thus, when integrated with Cu metal layers, improve Cu interconnect performance (about 30%) and lifetime (by orders of magnitude). The advantage of SWCNT bundle vias in controlling the back-end temperature, as shown in this paper, will also have significant implications for emerging technologies such as 3-D ICs where thermal management is a primary concern.

All these advantages, and the corresponding requirements from a process technology perspective that will be needed to make SWCNT bundles a viable interconnect material, are summarized in Table II.

### Table II

<table>
<thead>
<tr>
<th></th>
<th>Local</th>
<th>Intermediate</th>
<th>Global</th>
<th>Via/Contact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance</td>
<td>Same as Cu</td>
<td>~50% lower delay</td>
<td>30-40% lower delay</td>
<td>&lt;1% delay penalty</td>
</tr>
<tr>
<td>Reliability</td>
<td>Exceeds Cu</td>
<td>Exceeds Cu</td>
<td>Exceeds Cu</td>
<td>Exceeds Cu</td>
</tr>
<tr>
<td>Process Requirements</td>
<td>Dense bundle; Low contact resistance</td>
<td>Dense bundle</td>
<td>Dense bundle</td>
<td>Dense bundle; Low contact resistance</td>
</tr>
</tbody>
</table>

REFERENCES


Hong Li (S’07) received his B.E. degree in electronic engineering from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2003 and his M.S. degree from the Department of Electronic Engineering at Shanghai Jiao Tong University, Shanghai, China, in 2008. Since 2007 he has been working toward the Ph.D degree in Prof. Banerjee’s Nanoelectronics Research Laboratory in the Electrical and Computer Engineering department at the University of California, Santa Barbara.

Since December 2008, he has been the lead graduate student working on the development of a state-of-the-art carbon nanomaterial fabrication facility at UCSB’s California NanoSystems Institute (CNSI). His research interests are focused on the modeling and design of carbon nanomaterials for electronic applications—especially VLSI interconnects and passive devices. Mr. Li was awarded a number of prizes for outstanding academic performance as an undergraduate, including an Excellent Bachelor’s Thesis Award.

Franz Kreupl (M’08) received the Diploma in physics and the Dr. rer. nat. degree in physics from the University of Regensburg, Regensburg, Germany, in 1994 and 1998, respectively.

In 1998, he joined the Central Research Department of Siemens, Munich, Germany, where he was involved in interconnect technology and high-density integration of ferroelectric and magnetic memories. From 2000 to 2005, he was the Project Manager for carbon nanotubes, nanowires, and new materials in Infineon’s Corporate Research Department. From 2006 to 2009, he was with the Technology Innovation Department, Qimonda, Neubiberg, Germany, where he was engaged in emerging memories and memory systems. Since May 2009, he has been a Principal Engineer with the SanDisk Corporation in Milpitas, California. He has developed the first integrated carbon nanotube vias (2001), the first precise localized growth of single nanotubes (2002), the first nanotube power transistor (2003), the shortest nanotube transistor (2004), and the first application of graphene layers in dynamic RAMs (DRAMs) (2005). He has authored or coauthored more than 50 publications, contributed three book chapters, holds 50 patents, and has filed more than 70 patent families.

Dr. Kreupl received the Innovation Award in 2008 for developing the first application of graphene layers in DRAMs. He has given more than 70 invited talks. He was a member of the technical program committees of the IEEE International Conference on Nanotechnology (NANO), the European Solid-State Device Research Conference (ESSDERC), and the IEEE International Electron Devices Meeting (IEDM). He is a member of the International Technology Roadmap for Semiconductors (ITRS) Working Group on Emerging Research Devices.

Kaustav Banerjee (S’92–M’99–SM’03) received the Ph.D. degree in electrical engineering and computer sciences from the University of California, Berkeley, in 1999.

He has held summer/visiting positions at Texas Instruments Incorporated, Dallas, from 1993 to 1997, and the Swiss Federal Institute of Technology, Lausanne, Switzerland, in 2001. From 1999 to 2001, he was a Research Associate with the Center for Integrated Systems, Stanford University, Stanford, CA. From February to August 2002, he was a Visiting Faculty with the Circuit Research Laboratories, Intel, Hillsboro, OR. In July 2002, he joined the Faculty of the Department of Electrical and Computer Engineering, University of California, Santa Barbara, where he has been a Full Professor since 2007. His current research interests include nanometer-scale issues in high-performance/low-power VLSI as well as circuits and systems issues in emerging nanoelectronics. His research has been chronicled in over 160 journal and refereed international conference papers and two book chapters on 3-D ICs. He has also coedited the book Emerging Nanoelectronics: Life With and After CMOS (Springer-Verlag, 2004).

Prof. Banerjee was the recipient of a number of awards in recognition of his work including the Association for Computing Machinery (ACM) Special Interest Group on Design Automation (SIGDA) Outstanding New Faculty Award in 2004, the Best Paper Award at the Design Automation Conference in 2001, the IEEE Micro Top Picks Award in 2006, and the IBM Faculty Award in 2008. He was a member of the technical program committees of several leading IEEE and ACM conferences, including the IEEE International Electron Devices Meeting (IEDM), the Design Automation Conference (DAC), the International Conference on Computer Aided Design (ICCAD), and the IEEE International Reliability Physics Symposium (IRPS). From 2005 to 2008, he was a member of the Nanotechnology Committee of the IEEE Electron Devices Society. He was also a member of the organizing committee of the IEEE International Symposium on Quality Electronic Design (ISQED), at various positions including the Technical Program Chair in 2002 and the General Chair in 2005. He is a Distinguished Lecturer of the IEEE Electron Devices Society.