Table II summarizes the results of ASIC implementation along with other LDPC decoders available in the literature [7], [9]–[11]. For fair comparison with other reported architectures, throughput results for the proposed architecture were obtained assuming five iterations, 96 CNPUs, rate 1/2 code, and block length of 2304 bits. However, with an increase in the number of iterations the throughput would reduce, for example, with 15 iterations the throughput would be 364 Mbps. When operating at smaller code lengths, all the 96 CNPUs are not needed and hence can be shutdown to save power. Throughput can be further increased by adding additional pipeline stages but increases the overhead on area and power dissipation. For example, by adding one pipeline stage in the min unit the LDPC decoder can be run at a maximum clock frequency of 200 MHz. However, the area and power dissipation would be 1.3 million gates and 2 W respectively. The energy per bit per iteration of the proposed architecture is 87.10 pJ/bits/iteration. Although [7] has energy per bit per iteration of 7.4 pJ/bits/iteration, it only supports single code rate and a small code length regular LDPC code. Furthermore, the proposed architecture supports any check node degree, without the requirement of any reconfiguration and can accommodate changes to code rate and code length on-the-fly.

VI. Conclusion

A generic scalable architecture for the min function involved in the *min-sum/offset-min-sum* algorithm is proposed that supports both irregular and regular LDPC codes with check nodes of any degree. For an LDPC code with check node of maximum degree d, the proposed architecture consists of $2(d-2)$ 2×1 multiplexers and $3(d-2)$ two-input CSUs. The unpipelined version of proposed architecture has latency of $[2\lceil \log_2(d) \rceil - 2]t_{dc}$ when $\lceil \log_2(d) \rceil - \log_2(d) < \log_2(4/3)$ else $[2\lceil \log_2(d) \rceil - 3]t_{dc}$, with t_{dc} representing the delay of a two-input CSU. It also fares well against LDPC decoders currently available in the literature in terms of area, throughput, code rates/lengths supported, reconfigurability, and scalability. By performing the minimum search in a parallel fashion, the proposed min unit removes the bottleneck toward achieving a higher throughput generic scalable LDPC decoder architecture.

References

[1] R. G. Gallager, *Low Density Parity Check Codes*. Cambridge, MA: MIT Press, 1963.

[2] D. J. C. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 399–431, Mar. 1999.

[3] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, "Design of capacity-approaching low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 619–637, Feb. 2001.

[4] J. Chen and M. P. C. Fossorier, "Near optimum universal belief propagation based decoding of low-density parity check codes," *IEEE Trans. Commun.*, vol. 50, no. 3, pp. 406–414, Mar. 2002.

[5] J. Chen, A. Dholakia, E. Eleftheriou, M. P. Fossorier, and X.-Y. Hu, "Reduced-complexity decoding of LDPC codes," *IEEE Trans. Commun.*, vol. 53, no. 8, pp. 1288–1299, Aug. 2005.

[6] M. Karkotti and J. R. Cavallaro, "Semi-parallel reconfigurable architectures for real-time LDPC decoding," in *Int. Conf. Inform. Technol. (ITCC)*, 2004, vol. 1, pp. 579–585.

[7] A. Darabiha, A. C. Carusone, and F. R. Kschischang, "Power reduction techniques for LDPC decoders," *IEEE J. Solid-State Circuits*, vol. 43, no. 8, pp. 1835–1845, Aug. 2008.

[8] M. Bimberg, M. B. S. Tavares, E. Matúš, and G. P. Fettweis, "A high-throughput programmable decoder for LDPC convolutional codes," in *Proc. IEEE Int. Conf. Application-Specific Syst., Arch., and Process. (ASAP)*, Jul. 2007, pp. 239–246.

[9] M. Karkooti, P. Radosavljevic, and J. R. Cavallaro, "Configurable, high throughput, irregular LDPC decoder architecture: Tradeoff analysis and implementation," in *Proc. IEEE Int. Conf. Application-Specific Syst., Arch., and Process. (ASAP)*, 2006, pp. 360–367.

[10] T. Brack, F. Kienle, and N. Wehn, "Disclosing the LDPC code decoder design space," in *Proc. Design, Automation and Test in Europe (DATE)*, Mar. 2006, p. 6.

[11] K. Gunnam, G. Choi, W. Wang, and M. Yeary, "Multi-rate layered decoder architecture for block LDPC codes of IEEE 802.11n wireless standard," in *Proc. IEEE Int. Symp. Circuits and Syst. (ISCAS)*, May 2007, pp. 1645–1648.

[12] Y. Sun, M. Karkotti, and J. R. Cavallaro, "High throughput, parallel, scalable, LDPC encoder/decoder architecture for OFDM systems," in *Proc. IEEE Dallas/CAS Workshop*, Oct. 2006, pp. 39–42.

[13] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.

Evaluating Carbon Nanotube Global Interconnects for Chip Multiprocessor Applications

Sudeep Pasricha, Fadi J. Kurdahi, and Nikil Dutt

Abstract—In ultra-deep submicrometer (UDSM) technologies, the current paradigm of using copper (Cu) interconnects for on-chip global communication is rapidly becoming a serious performance bottleneck. In this paper, we perform a system level evaluation of Carbon Nanotube (CNT) interconnect alternatives that may replace conventional Cu interconnects. Our analysis explores the impact of using CNT global interconnects on the performance and energy consumption of several multi-core chip multiprocessor (CMP) applications. Results from our analysis indicate that with improvements in fabrication technology, CNT-based global interconnects can significantly outperform Cu-based global interconnects.

Index Terms—Carbon nanotubes, chip multiprocessors, global interconnects, system level design.

I. Introduction

As technology scales into the ultra-deep submicrometer (UDSM) region, poor interconnect scaling has made interconnect design a major bottleneck for emerging multi-core chip multiprocessors (CMPs) that have tens to hundreds of components integrated on a single chip [1]. According to the International Roadmap for Semiconductors (ITRS) [2], global interconnect performance in particular has become one of the semiconductor industry's topmost challenges. With technology scaling, conventional copper (Cu)-based global interconnects have become increasingly susceptible to electromigration, rising crosstalk coupling noise, and parasitic resistivity, which has reduced reliability and performance [3]. There is therefore a critical need to investigate innovative global interconnect alternatives to Cu.

Recently, there has been tremendous interest in carbon nanotubes (CNTs) as a possible replacement for Cu interconnects in future technologies [4], [5]. Depending on the direction in which they are rolled (referred to as *chirality*) CNTs can behave either as semiconductors or conductors. Conducting (or metallic) CNTs are highly resistant to electromigration and other sources of physical breakdown, can support very high current densities with very little performance degradation, and have much better conductivity properties compared to Cu [4].

CNTs can be broadly classified into single-walled carbon nanotubes (SWCNTs) [6], [7] and multi-walled carbon nanotubes (MWCNTs) [8], [9]. SWCNTs consist of a single sheet of graphene rolled into a cylindrical tube, with a diameter in the nanometer range. MWCNTs consist of two or more SWCNTs concentrically wrapped around each other, with diameters ranging from a few to several hundred nanometers. For on-chip global interconnects, bundles of SWCNTs and mixed SWCNT/MWCNTs are also of special interest because of their superior conductivity properties. A SWCNT bundle consists of several SWCNTs packed in parallel [8], whereas a mixed SWCNT/MWCNT bundle consists of a combination of SWCNTs and MWCNTs packed in parallel [9].

The question arises: *how do CNTs compare against Cu as global interconnect materials?* While there has been a lot of interest in CNT-based interconnects in recent years, there are few studies that compare CNT interconnects with Cu interconnects at the system-level for emerging CMPs. In this paper, we present such a comparative analysis. We investigate the performance and energy impact of using the four CNT-based global interconnect alternatives discussed above in place of conventional Cu global interconnects at the system-level for several CMP applications. Detailed RLC equivalent circuit models are used to capture the statistical distribution of metallic CNTs while accurately incorporating recent experimental and theoretical results on inductance, crosstalk capacitance and ohmic resistance. These models are incorporated into a system level simulation environment to estimate the impact of using CNT-based global interconnects on the overall performance and energy consumption of several heterogeneous CMP applications with diverse data traffic profiles.

II. RELATED WORK

Recently, there has been a lot of interest in studying the properties of CNTs. Researchers have developed *RLC* circuit models for CNT interconnect alternatives and compared their performance with Cu interconnects at the circuit level. Raychowdhury *et al.* [6] and Naeemi *et al.* [7] presented *RLC* circuit models for isolated SWCNTs. Srivastava *et al.* [10] presented circuit models for SWCNT bundles and compared sparse and dense bundle propagation delays with propagation delays in conventional Cu interconnects. Naeemi *et al.* [8] presented an *RLC* circuit model for MWCNTs which was modified by Li *et al.* [11] for more accurate conductance analysis. Nieuwoudt *et al.* [12] explored the problem of optimal MWCNT dimensions and diameter sizing. Haruehanroengra *et al.* [9] and Wang *et al.* [13] proposed conductance and inductance models for mixed SWCNT/MWCNT bundles.

To date however, little work has been done to analyze the impact of using CNT interconnects instead of Cu interconnects at the system-level. The work that comes closest to our goal is by Rossi *et al.* [14] in which a dual-walled CNT global bus was presented and experiments performed to show that it has a lower propagation delay compared to SWCNT buses. Unlike [14], in this paper we explore a variety of CNT based global interconnects for emerging CMP applications and present comparisons with conventional Cu global interconnects.

III. CNT GLOBAL INTERCONNECT *RLC* MODELS

In this section we present a brief overview of *RLC* circuit models for the four CNT global interconnect candidates.

1) SWCNT: Each conducting channel in a CNT has two modes due to spin degeneracy, and has a quantum resistance $R_Q = h/2e^2 = 12.9$ KΩ, kinetic inductance $L_K = h/4e^2 v_F = 8$ nH/μm, and quantum capacitance $C_Q = 4e^2/hv_F = 194$ aF/μm, where h is the Planck's constant, e is the elementary charge, and V_F is the Fermi-velocity of CNTs [5], [15]. For low bias voltage (<0.16 V for global wires [16]) there also exists scattering induced resistance per unit length $R_S = R_Q/\lambda$, where λ is the mean free path (MFP) of CNT ($\lambda \approx 1$ μm for SWCNTs [6]). A SWCNT has two channels due to lattice degeneracy, hence a total of four conduction modes. For parallel SWCNTs, there also exists an electrostatic capacitance per unit length with the ground $C_{EG} = 2\pi\varepsilon/\cos h^{-1}(2H/d_t)$ and a coupling capacitance per unit length with any adjacent SWCNTs $C_{EG} = \pi\varepsilon/\cos h^{-1}(s/d_t)$, where ε is the dielectric permittivity, H is the distance between the SWCNT and ground, s is the inter-SWCNT spacing, and d_t is the SWCNT diameter. The quantum capacitance is in series with a parallel combination of electrostatic and coupling capacitances.

2) SWCNT Bundle: Since an SWCNT bundle consists of several individual SWCNTs in parallel, the effective resistance $R_{\text{SWCNT\_bundle}} = R_{\text{SWCNT}}/n_{\text{CNT}}$ where n_{CNT} is the number of conducting SWCNTs in the bundle, characterized by the metallic density (P_m) parameter. P_m refers to the probability that an SWCNT in the bundle is metallic (i.e., conducting). The value of $P_m = 1/3$ [1] with today's best fabrication techniques. The coupling and electrostatic capacitances to the ground of a highly metallic SWCNT bundle ($P_m = 1$) have been analyzed extensively using the field solvers RAPHAEL [6] and FastCap [16], and found to be equal to the respective capacitances of a Cu wire with the same cross-sectional dimensions. For more realistic SWCNT metallic densities ($P_m = 1/3$), capacitance was shown to decrease very slowly ($\sim$20% lower even for as few as four SWCNTs at the corners), than for a densely packed bundle [6]. The effective quantum capacitance of a SWCNT bundle is negligible compared to its electrostatic counterparts [16]. The inductance of the SWCNT bundle can be derived using the partial element equivalent circuit (PEEC) model [17] as $L_{\text{SWCNT\_bundle}} = (L_K/4n_{\text{CNT}} + L_M).l$, where magnetic inductance $L_m \approx 1.6$ pH/μm at the 22 nm node.

3) MWCNT: The building blocks of MWCNTs are concentric SWCNTs of varying diameters, with the number of shells $N_S = 1 + (D_{\text{outer}}D_{\text{inner}})/2\delta$, where $\delta = 0.34$ nm (van der Waals distance) is the spacing between adjacent concentric shells, and D_{outer} and D_{inner} are the maximum and minimum shell diameters. The ratio of $D_{\text{inner}}/D_{\text{outer}}$ has been observed to vary from 0.35 to 0.8 [8]. The approximate number of conduction channels of each shell in a MWCNT is [9]: $N_{\text{chan/shell}} \approx (aD+b)$, where D is the shell diameter (> 6 nm), $a = 0.0612$ nm^{-1}, and $b = 0.425$. Similar to a SWCNT, for each shell, resistance $R_{\text{shell}} = (R_Q + R_S.l)/N_{\text{chan/shells}}$, quantum capacitance $C_{Q/\text{shell}} = C_Q.N_{\text{chan/shell}}$ and kinetic inductance $L_{K/\text{shell}} = L_K/N_{\text{chan/shell}}$. An inter-shell resistance ($R_r \approx 10$ kΩ/μm [14]) must also be considered to account for the inter-shell tunnel transport phenomenon. The metallic shell in an MWCNT constitutes an effective electrostatic shield for its inner shells. Thus the capacitance between an internal shell and ground, and between nonadjacent shells can be safely neglected. The electrostatic capacitance per unit length between the outermost MWCNT shell and the ground and coupling capacitance between the outer shells of adjacent MWCNTs is calculated using the relations for SWCNTs, presented earlier. The electrostatic coupling capacitance between adjacent shells in a MWCNT (C_{ESC}) is derived from a conventional metallic coaxial configuration [18]: $C_{\text{ESC}} = 2\pi\varepsilon/ln(d1/d2)$, where $d1$, $d2$ are shell diameters. L_m is calculated using the equivalent conductivity method [13].

4) Mixed SWCNT/MWCNT Bundles: A mixed bundle consists of SWCNTs with a diameter d and MWCNTs with various diameters $D_{inner} \geq d_i \geq D_{outer}$. It has been shown [3] that the outer MWCNT diameters (D_{outer}) follow a normal (Gaussian) distribution. The resistance of a mixed bundle depends on the tube count, which in turn depends on the value of D_{outer}, the mean diameter $\overline{D_{outer}}$, and the standard deviation σD_{outer}. Details of resistance calculations can be found in [3] and [9]. The capacitive characteristics of a mixed bundle have been shown to predominantly be determined by the cross sectional dimensions of a bundle [1]. Therefore the capacitance is similar to that of a SWCNT bundle with the same dimensions. The total kinetic inductance of a mixed bundle is the parallel inductance value of all the conduction channels in the bundle, and calculated similar to that for a SWCNT bundle. L_m is calculated using the equivalent conductivity method [13].

IV. EXPERIMENTS

A. Propagation Delay Comparison

In our first experiment we compared the global wire delay of Cu and CNT interconnect alternatives across the 45–22 nm UDSM technology nodes. We performed electrical level simulation with HSPICE using the equivalent RLC circuit models for CNTs from the previous section and *RLC* models for Cu derived from [19] and [20]. We implemented optimal repeater sizing and insertion for both Cu and CNT wires [21]. The node driver resistance, load capacitance, and process parameters were obtained from ITRS specifications [2]. Since global wire width is much larger than minimum wire width (W_{min}) to improve delay, power and bandwidth characteristics, we used a global wire width value of $5W_{min}$ which optimizes the power-delay product for Cu wires [22]. An aspect ratio of three was chosen based on ITRS specifications [2]. For comparison purposes, we used the same width and aspect ratios for the SWCNT and mixed bundles. For the mixed bundle, we assumed $\overline{D_{outer}} = 4.2$ nm and $\sigma_{D_{outer}} = 1.25$ nm [3]. The SWCNT diameter was assumed to be 1 nm. The ratio between outer to inner diameters for MWCNTs was assumed to be 0.5. The spacing between adjacent wires was assumed to be the same as the wire width ($5W_{min}$). The CNT mean free path $\lambda = 1~\mu m$ and metallic density $P_m = 1/3$, while Cu MFP length was assumed to be 40 nm. All of these values are practically achievable today by using prevalent fabrication techniques.

Fig. 1 shows the average signal propagation speedup for CNTs. The y-axis depicts the ratio of propagation delay of copper $t(Cu)$ and CNTs $t(CNT)$, averaged for global wires (>1 mm in length) using Monte Carlo simulations across the 45–22 nm nodes. The isolated SWCNT has been omitted from the analysis and the figure because of its much higher propagation delay (almost $100\times$ more) compared to Cu, due to its extremely small cross-section area and higher resistance. Thus isolated SWCNTs are unsuitable as global interconnects. From the figure, all the other CNT alternatives have lower interconnect delays compared to Cu. The MWCNT wire in particular has a much lower delay due to its large concentric tubes having greater MFP lengths which improves conductivity over Cu. A mixed bundle has many large MWCNTs with several shells and more conduction channels than SWCNT bundles of the same dimensions. That is why mixed bundles are better conductors than SWCNT bundles.

B. System-Level Performance and Energy Analysis

For our next experiment, we selected several multi-core CMP applications to analyze the overall performance and energy impact of using CNT global interconnects at the system-level. The applications are selected from the well known SPLASH-2 benchmark suite [23], as well as from the networking domain. These applications are parallelized and implemented on multiple cores. The entire chip is assumed
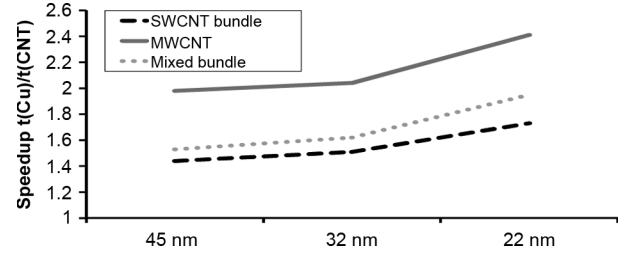

Fig. 1. Cu versus CNT global interconnect delay comparison (45–22 nm).

TABLE I
CMP APPLICATION IMPLEMENTATION CHARACTERISTICS

CMP applications	Description	cores	prog. processors	clusters	global links
Radix	Integer radix sort	18	4	3	10
Barnes	Galaxy evolution	26	6	4	18
Water-NSq	Forces/potentials of H_2O molecules	26	8	4	15
FFT	FFT kernel	28	6	4	12
Cholesky	Cholesky factorization kernel	33	7	5	19
Ocean	Ocean movements	35	10	5	24
Raytrace	3-D ray tracing	43	17	5	30
Netfilter	Packet processing and forwarding	49	22	6	32
Datahub	Packet processing and forwarding	68	26	8	34
SecurePck	Packet processing and forwarding	94	30	8	28

to be partitioned into clusters (or islands) of heterogeneous computation cores. Each cluster consists of tightly coupled cores (processors, memories etc.) optimized for dedicated tasks (e.g., packet encryption, image processing, etc.) and interconnected via local bus-based Cu links that support high data bandwidths. Global pipelined links (shared bus or point-to-point) are used to interconnect the clusters. Table I summarizes the implementation details of the CMP applications, which were modeled in SystemC [24] using a fast transaction-based bus cycle accurate (T-BCA) modeling abstraction [25]. The cores in each cluster were modeled at the behavioral level granularity, while the communication at the inter- and intra-cluster level was modeled at a cycle accurate granularity. Each of the applications was simulated with testbench traffic to quickly ($\sim$ few hours) and accurately estimate performance of the implementations. The various cores were interconnected using the AMBA AXI standard interface protocol, with a 32 bit address bus width and 64 bit read and write data buses.

Our analysis was performed for the 22-nm technology node (2016 node of ITRS [2]) and the interconnect fabric was clocked at frequencies ranging from 1–10 GHz to support data transfer rates in the hundreds of Gbps range for future high performance systems. A high level floorplanner (PARQUET [26]) was used to create an early layout of the CMP applications, and wire lengths were calculated using Manhattan distance. Latches were inserted on global interconnects wherever necessary, to pipeline interconnects and ensure correct synchronous operation.

Figs. 2(a) and (b) show the performance improvement (speedup in CMP application execution time) when using SWCNT bundle, mixed SWCNT/MWCNT and MWCNT bundle global interconnects, instead of Cu global interconnects, for the 22-nm node, with the interconnect fabric clocked at 10 GHz. Fig. 2(a) shows the results for the case when imperfect metal-CNT contacts are assumed. A high metal-CNT contact resistance of 100 KΩ is added at each metal-CNT junction to account for the excessive electron scattering at imperfect metal-CNT junctions. It can be seen that the application performance with CNT global interconnects is worse than with Cu global interconnects. This
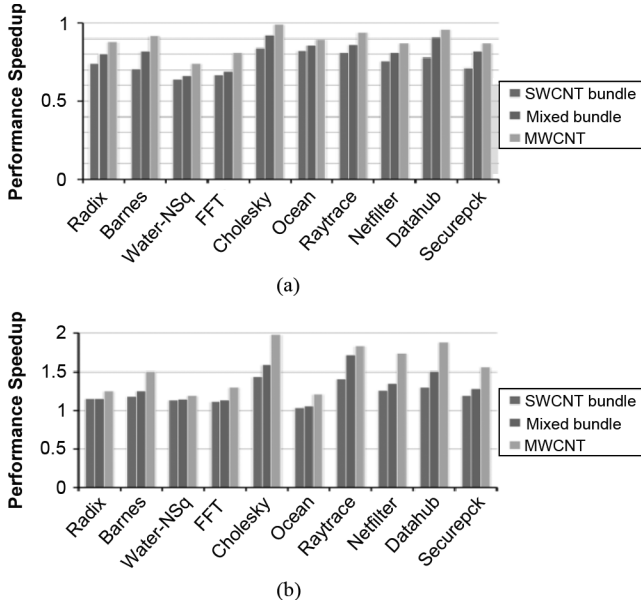
Fig. 2. System level performance speedup for CNTs global interconnects with (a) imperfect metal-CNT contacts and (b) perfect metal-CNT contacts.
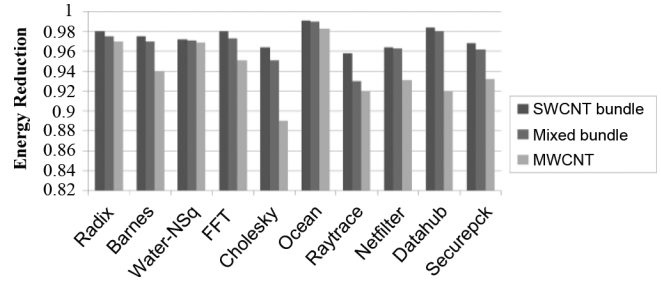

Fig. 3. Communication architecture energy consumption for CNT global interconnects, normalized to the energy dissipation with Cu base case.
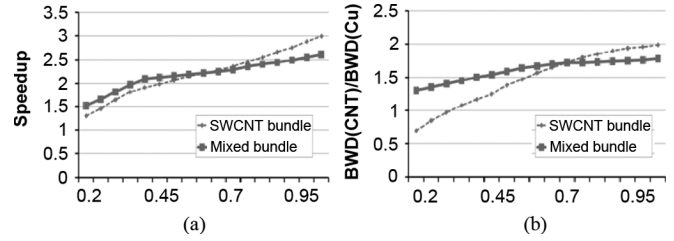

Fig. 4. (a) SWCNT bundle and mixed bundle. (a) Delay speedup over Cu for varying P_m. (b) Bandwidth density speedup over Cu for varying P_m.

is primarily an artifact of the poor metal-CNT contacts. However, many recent studies [27], [28] with state-of-the-art fabrication techniques have managed to reduce this contact resistance down to a very small value ($\sim 100\ \Omega$) for bundle and multi-walled CNTs. This has dramatically improved the viability of successful integration of CNTs with CMOS technology. Fig. 2(b) shows the application performance improvement for perfect contacts with negligible contact resistance ($\sim 100\ \Omega$). The applications now perform better with CNT global interconnects because of their lower signal propagation delays. The lower delays lead to more widely spaced and hence fewer pipeline latches on global inter-cluster interconnects, resulting in global data transfers taking fewer clock cycles. This improves overall application performance. Among the applications, *Ocean* and *Water-NSq* have a much smaller speedup compared to other applications, because of fewer inter-cluster data transfers, which reduces the advantage of having faster CNT global interconnects. On the other hand, the speedup for *Radix* and *FFT* is lower than other applications because of the smaller length of their global interconnects, which again reduces the impact of faster CNT global interconnects on application performance. Overall, MWCNT global buses provide the maximum speedup (up to 1.98×). SWCNT bundles and mixed SWCNT/MWCNT bundles achieve speedups of up to 1.3× and 1.5×, respectively.

Fig. 3 shows the change in communication architecture energy consumption (calculated using the methodology proposed by us in [29]) when CNT global interconnects are used instead of Cu global interconnects in the CMP applications. The communication architecture energy consumption shown accounts for energy consumed not just by switching on the wires but also due to switching and leakage in the latches and AXI interface logic. Since CNTs have a marginally lower capacitance and require fewer pipeline stages compared to Cu global interconnects, there is, in general, a reduction in energy consumption (1%–11%). The magnitude of this reduction depends on complex interactions between data patterns within each application.

At first glance, the modest performance and energy improvements over Cu interconnects may appear to be not so significant as to justify a migration to CNT-based interconnects in the future. However, in the next section, we present results that indicate more significant CNT gains with improvements in CNT fabrication technology.

C. Impact of MFP and CNT Metallic Density

Our experiments above assumed practical values for CNT metallic density ($P_m = 1/3$) and mean free path ($\lambda = 1\ \mu$m) while calculating performance and energy gains. The modest CNT bundle speedup obtained above can be improved if breakthroughs in fabrication technology in the future allow for greater metallic density (P_m) and longer MFP lengths (λ) for CNT bundles. Fig. 4(a) shows the interconnect delay speedup over Cu for SWCNT bundles and mixed bundles, with increasing values of metallic density P_m, from 0.3 (practical) to 1 (ideal), for a constant MFP $\lambda = 1\ \mu$m. It can be seen that there is a crossover point beyond which SWCNT bundle performance improves upon mixed bundle performance. This is because as the metallic density increases, the mixed bundle has fewer metallic nanotubes than the SWCNT bundle due to (relatively) large MWCNTs which prevent a tight packing. The SWCNT bundle has more tightly packed and thus more metallic tubes which improve conductivity and performance.

Fig. 4(b) shows a similar trend when bandwidth density of CNT bundle interconnects is compared with Cu, for varying metallic densities. It is clear from the figure that SWCNT bundle bandwidth density is superior to that of Cu for metallic densities greater than 0.65. SWCNT bundle bandwidth density also improves upon mixed bundle bandwidth density for higher metallic densities, due to more tightly packed metallic tubes. Figs. 5 and 6 show the combined performance improvement for various configurations of metallic density (P_m) and MFP (λ) lengths for the SWCNT bundle and mixed SWCNT/MWCNT bundle interconnects, respectively.

As far as energy consumption is concerned, with increasing metallic density, the capacitance of CNT bundle interconnects increases somewhat and approaches that of Cu interconnects. However, as the CNT propagation delay is reduced with the increase in MFP length and metallic density, the number of pipeline stages in long interconnects is reduced, leading to a reduction in power dissipation and energy consumption. Our experiments with the CMP applications indicated a potential communication architecture energy saving of as much as 30% with the advances in fabrication technology discussed in this section.
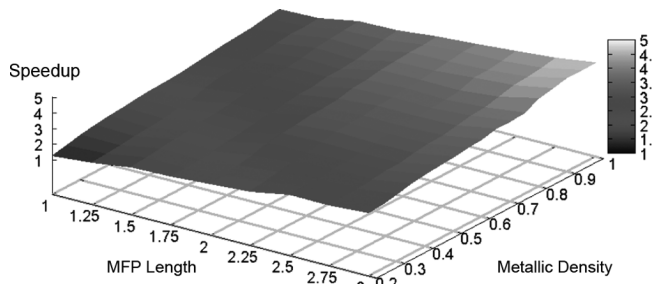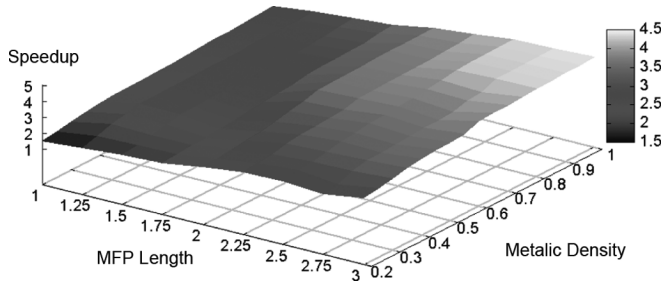
Fig. 5. SWCNT bundle speedup over Cu.

Fig. 6. Mixed bundle speedup over Cu.

It is clear from the results in this section that increasing MFP length and metallic densities can lead to a substantial improvement in delay (and hence performance) speedup for CNT bundle interconnects—as much as $4.9\times$ for SWCNT bundles and up to $4.4\times$ for mixed SWCNT/ MWCNT bundles. Simultaneously, energy consumption in the communication fabric can also be notably reduced. Improving CNT fabrication technology in order to achieve such performance and energy gains is an area of tremendous research activity today [30], [31]. Already, promising recent work has made a strong case for using arrays of CNT interconnects integrated on top of silicon CMOS ICs [32].

V. CONCLUSION

In this paper, we presented a comparative analysis of the performance and energy impact of using CNT-based global interconnects over Cu for heterogeneous multi-core CMP applications. The experimental results indicate that while SWCNTs are not as suitable for global interconnect buses due to their large ohmic resistance and very high delays, global MWCNT buses can provide performance speedups of up to $1.98\times$ for CMP applications. Global interconnect buses implemented with SWCNT bundles and mixed bundles also lead to performance gains over Cu global buses of up to $1.3\times$ and $1.5\times$, respectively. These gains can be significantly improved if the CNT MFP lengths and metallic densities are increased with advances in fabrication technology that are actively being explored today.

REFERENCES

[1] S. Pasricha and N. Dutt, *On-Chip Communication Architectures*. Burlington, MA: Morgan Kauffman, 2008.

[2] ITRS, "International Technology Roadmap for Semiconductors," 2005.

[3] J. Li, Q. Ye, A. Cassell, H. T. Ng, R. Stevens, J. Han, and M. Meyyappan, "Bottom-up approach for carbon nanotube interconnects," *Appl. Phys. Lett.*, vol. 82, no. 15, pp. 2491–2493, Apr. 2003.

[4] M. S. Dresselhaus, G. Dresselhaus, and A. Jorio, *Carbon Nanotubes: Synthesis, Structure, Properties, and Applications*. New York: Springer, 2001.

[5] P. J. Burke, "Luttinger liquid theory as a model of the gigahertz electrical properties of carbon nanotubes," *Trans. NANO*, vol. 1, no. 3, pp. 129–144, Sep. 2002.

[6] A. Raychowdhury and K. Roy, "Modeling of metallic carbon nanotube interconnects for circuit simulation and a comparison with Cu interconnects for scaled technologies," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 25, no. 1, pp. 1361–1364, Jan. 2006.

[7] A. Naeemi and J. D. Meindl, "Design and performance modeling for single-walled carbon nanotubes as local, semiglobal, and global interconnects in gigascale integrated systems," *IEEE Trans. Electron Device Lett.*, vol. 54, no. 1, pp. 26–37, Jan. 2007.

[8] A. Neaeemi and J. D. Meindl, "Compact physical models for multiwall carbon-nanotube interconnects," *IEEE Trans. Electron Device Lett.*, vol. 27, no. 5, pp. 338–340, May 2006.

[9] S. Haruehanroengra and W. Wang, "Analyzing conductance of mixed carbon-nanotube bundles for interconnect applications," *IEEE Trans. Electron Device Lett.*, vol. 28, no. 8, pp. 756–759, Aug. 2007.

[10] N. Srivastava and K. Banerjee, "Performance analysis of carbon nanotube interconnects for VLSI applications," in *Proc. ICCAD*, 2005, pp. 383–390.

[11] H. Li, W.-Y. Yin, K. Banerjee, and J.-F. MaoCircuit, "Modeling and performance analysis of multi-walled carbon nanotube interconnects," *IEEE Trans. Electron Devices*, vol. 55, no. 6, pp. 1328–1337, Jun. 2008.

[12] A. Niewoudt and Y. Massoud, "Investigating the design, performance, and reliability of multi-walled carbon nanotube interconnect," in *Proc. ISQED*, 2008, pp. 691–696.

[13] W. Wang, S. Haruehanroengra, L. Shang, and M. Liu, "Inductance of mixed carbon nanotube bundles," *Micro & Nano Lett., IET*, vol. 2, no. 2, pp. 35–39, Jun. 2007.

[14] D. Rossi, J. M. Cazeaux, C. Metra, and F. Lombardi, "Modeling crosstalk effects in CNT bus architectures," *IEEE Trans. Nanotechnol.*, vol. 6, no. 2, pp. 133–145, Mar. 2007.

[15] K. Banerjee, H. Li, and N. Srivastava, "Current status and future perspectives of carbon nanotube interconnects," in *Proc. NANO*, 2008, pp. 432–436.

[16] H. Cho, K-H. Koo, P. Kapur, and K. C. Saraswat, "Modeling of the performance of carbon nanotube bundle, cu/low-k and optical on-chip global interconnects," in *Proc. SLIP*, 2007, pp. 81–88.

[17] M. W. Beattie and L. T. Pileggi, "Inductance 101: Modeling and extraction," in *Proc. DAC*, 2001, pp. 323–328.

[18] E. R. Dobbs, *Basic Electromagnetism*. London, U.K.: Chapman Hall, 1993.

[19] A. F. Mayadas and M. Shatzkes, "Electrical-resistivity model for polycrystalline films: The case of arbitrary reflection at external surfaces," *Phys. Rev. B*, vol. 1, no. 4, pp. 1382–1389, 1970.

[20] E. H. Sondheimer, "The mean free path of electrons in metals," *Adv. Phys.*, vol. 1, no. 1, pp. 499–537, 1952.

[21] Y. I. Ismail and E. G. Friedman, "Effects of inductance on the propagation delay and repeater insertion in VLSI circuits," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 8, no. 2, pp. 195–206, Apr. 2000.

[22] M. A. El-Moursy and E. G. Friedman, "Optimum wire sizing of RLC interconnect with repeaters," in *Proc. GLSVLSI*, 2003, pp. 27–32.

[23] S. C. Woo, M. Ohara, E. Torrie, J. P. Singh, and A. Gupta, "The SPLASH-2 programs: Characterization and methodological considerations," in *Proc. ISCAS*, 1995, pp. 1–13.

[24] Open SystemC Initiative (OSCI), "SystemC Initiative," [Online]. Available: www.systemc.org

[25] S. Pasricha, N. Dutt, and M. Ben-Romdhane, "Extending the transaction level modeling approach for fast communication architecture exploration," in *Proc. DAC*, 2004, pp. 113–118.

[26] S. N. Adya and I. L. Markov, "Fixed-outline floorplanning: Enabling hierarchical design," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 11, no. 6, pp. 1120–1135, Dec. 2003.

[27] O. Hjortstam, P. Isberg, S. Söderholm, and H. Dai, "Can we achieve ultra-low resistivity in carbon nanotube-based metal composites?," *App. Phy. MSP*, vol. 78, no. 8, pp. 1175–1179, 2004.

[28] Z. Liu, L. Ci, N. Bajwa, P. M. Ajayan, and J.-Q. Lu, "Benchmarking of metal-to-carbon nanotube side contact resistance," in *Proc. IITC*, 2008, pp. 144–146.

[29] S. Pasricha, Y. Park, F. Kurdahi, and N. Dutt, "System-level power-performance trade-offs in bus matrix communication architecture synthesis," in *Proc.* CODES + ISSS, Oct. 2006, pp. 300–305.

[30] Liu, N. Bajwa, L. Ci, S. H. Lee, S. Kar, P. M. Ajayan, and J.-Q. Lu, "Densification of carbon nanotube bundles for interconnect application," in *Proc. IITC*, 2007, pp. 201–203.

[31] Z. Zhu, K.-S. Moon, B. Bertram, D. W. Hess, and C. P. Wong, "Assembling carbon nanotube bundles using transfer process for fine-pitch electrical interconnect applications," in *Proc. EETC*, 2007, pp. 1981–1985.

[32] G. F. Close, S. Yasuda, B. Paul, S. Fujita, and H.-S. P. Wong, "A 1 GHz integrated circuit with carbon nanotube interconnects and silicon transistors," *Nano Lett.*, vol. 8, no. 2, pp. 706–709, 2008.