REPORT

DEVICE TECHNOLOGY

Carbon nanotube transistors scaled to a 40-nanometer footprint

Qing Cao,* Jerry Tersoff, Damon B. Farmer, Yu Zhu, Shu-Jen Han

The International Technology Roadmap for Semiconductors challenges the device research community to reduce the transistor footprint containing all components to 40 nanometers within the next decade. We report on a p-channel transistor scaled to such an extremely small dimension. Built on one semiconducting carbon nanotube, it occupies less than half the space of leading silicon technologies, while delivering a significantly higher pitch-normalized current density—above 0.9 milliampere per micrometer at a low supply voltage of 0.5 volts with a subthreshold swing of 85 millivolts per decade. Furthermore, we show transistors with the same small footprint built on actual high-density arrays of such nanotubes that deliver higher current than that of the best-competing silicon devices under the same overdrive, without any normalization. We achieve this using low-resistance end-bonded contacts, a high-purity semiconducting carbon nanotube source, and self-assembly to pack nanotubes into full surface-coverage aligned arrays.

The International Technology Roadmap for Semiconductors (ITRS) shared across chip manufacturers, materials suppliers, and apparatus makers (*1*) has guided the miniaturization of logic transistors by the progress of technology "nodes," with smaller numbers indicating newer technologies for smaller and faster devices. Silicon field-effect transistors (Si FETs) are currently being produced at the 14-nm node, which has an overall lateral footprint of about 90 to 100 nm (*2*), but they are already near their scaling limits. The ITRS roadmap projects that the device footprint will reach 40 nm a decade from now with the 3-nm node, where both the device gate length (L_g) and contact length (L_c) are reduced to ~10 nm with ~5-nm-wide spacers separating each side of the gate from source-drain electrodes (*1*).

Alternative technologies are being aggressively explored for the next digital switch to fulfill these stringent requirements. Transistors based on semiconducting carbon nanotubes (s-CNTs) are considered to be among the most promising candidates (*3*). The intrinsic thinness of s-CNTs (about 1 nm in diameter) enables superb electrostatic control to minimize passive power dissipation in the off state, and their high saturation velocity for electrons and holes allows devices to switch at a given frequency under a much lower drive voltage (V_{DD}), which reduces dynamic power consumption. Intensive research efforts from both academia and industry in the past 20 years have fueled the remarkable development of nanoelectronics based on carbon nanotubes, resulting in important demonstrations, including devices with ideal gate-all-around geometry (*4*), complementary logic using standard semiconductor processes (*5*), mass production of more than 10,000 individual s-CNT transistors (*6*), and functional nanotube circuits up to a primitive microprocessor (*7*).

These past demonstrations represent important technology milestones but were invariably carried out on devices much larger than current Si FETs without scaling the entire nanotube transistor to the targeted 40-nm footprint. To evaluate the scalability of s-CNT transistors in experiment, great progress has been made in reducing one key device component, L_g, to 10 and even 5 nm (*8*, *9*). Still, those previous high-performance sub-10 nm L_g nanotube devices (*8*, *9*) employed 100- to 200-nm-long source and drain contacts overlapping with the s-CNT to achieve low contact resistance (*10*). Connecting the open ends of s-CNTs to molybdenum contacts could permit the reduction of L_c down to 10 nm without increasing resistance (*11*). However, that process requires annealing at 850°C, which is incompatible with the fabrication of channels below 60 nm due to the lack of the required structural stability for even refractory metals such as molybdenum at such high temperatures (*11*). Pure cobalt or nickel film might form end-bonded contacts to sparse individual nanotubes at 400° to 600°C through carbon

IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598, USA.
*Corresponding author. Email: qcao@us.ibm.com

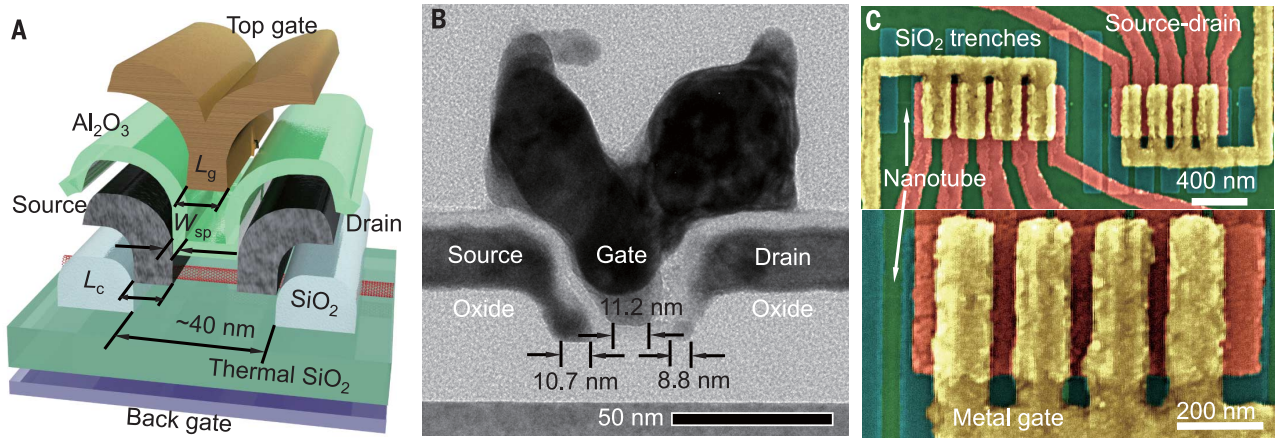

Fig. 1. Illustration and electron microscopy images of extremely scaled individual s-CNT transistors. (**A**) Schematic exploded view showing the oxide trench defining the 40-nm device footprint, the end-bonded source-drain contacts to the s-CNT channel, the 5-nm Al_2O_3 gate dielectric and spacer, and the top-gate device structure. L_c, W_{sp}, and L_g mark the device contact length, spacer width, and gate length, respectively. (**B**) Cross-sectional TEM image of a device as in (A) through the gate illustrating the profile of the oxide trench, the contacts, the dielectric, and the gate electrode to ensure the accuracy in defining L_c, W_{sp}, and L_g, as well as the overall device footprint. (**C**) SEM images of two sets of transistors made on a single s-CNT with their overall device footprint of 40 nm (magnified view also shown in the bottom frame) and 85 nm as defined by the space of SiO_2 trenches. The metal gate (gold) and the source-drain (dark red) electrodes, the oxide bars (dark cyan), and the substrate with the s-CNT (green) are colorized to highlight each component.

dissolution (*12*). However, their much lower melting temperature limited the smallest attainable L_c to 30 nm and L_g to 60 nm because of poor structural stability even at these temperatures (*12*). No process has been reported with even the potential to satisfy the extremely scaled 40-nm footprint. Moreover, these past device dimension scalability studies only focused on transistors built on individual nanotubes. A practical technology requires that arrays of s-CNTs perform together to provide adequate drive current. Thus, the ultimate performance requirements are far beyond anything that has been demonstrated to date.

We report a complete high-performance p-channel s-CNT transistor scaled to the 40-nm footprint, as required by the 3-nm technology node and beyond. We used a top-gate structure together with end-bonded contacts formed with a low-temperature process (650°C). We compare the pitch-normalized performance superiority of single s-CNT transistors at this dimension against state-of-the-art silicon technologies. Moreover, we demonstrate the fabrication of more technology-relevant high-performance nanotube-array devices with the same footprint, using a high-purity s-CNT source, self-assembly to pack nanotubes into full surface-coverage aligned arrays, and low-resistance end-bonded contacts. These s-CNT–array transistors exhibited a high saturation on-state current above 1.2 mA μm^{-1} and conductance above 2 mS μm^{-1}, which exceeds that of the best-competing silicon devices when they are benchmarked under the same gate overdrive and source-drain bias (V_{DS}), without any normalization.

A schematic structure of our transistor is shown in Fig. 1A, and the actual device is shown in a cross-sectional transmission electron microscopy (TEM) micrograph in Fig. 1B. The detailed fabrication process is described in the supplementary materials and in fig. S1 (*13*). The overall footprint of the transistor is confined by the space of a SiO_2 trench, in a way similar to the distance between edges of two neighboring contact vias in Si FETs, to only 40 nm as measured from the bottom of the trench where the ~1-nm-thick s-CNT resides and connects with the source-drain contacts. Electrodes extending on top of that can be considered as part of the local interconnects, whose presence does not affect the assessment of device scalability at the individual transistor level. A single s-CNT, sitting on top of a 20-nm-thick thermal oxide grown on a silicon handle wafer, was end-bonded to 10-nm-wide source and drain Co-Mo alloy contacts by a solid-state carbide-formation reaction.

The adoption of end-bonded contacts allowed further reduction of L_c to make even smaller transistors without compromising performance. The cobalt here apparently acted as a catalyst, reducing the reaction temperature by 200°C compared with pure Mo (*11*). Such greatly reduced reaction temperature allowed end contacts to form while maintaining the structural integrity of sub-20-nm gaps between source and drain electrodes [see supplementary text 1 and fig. S2 (*13*)]. A 5-nm-thick Al_2O_3 film deposited conformally on top by atomic layer deposition (ALD) defined

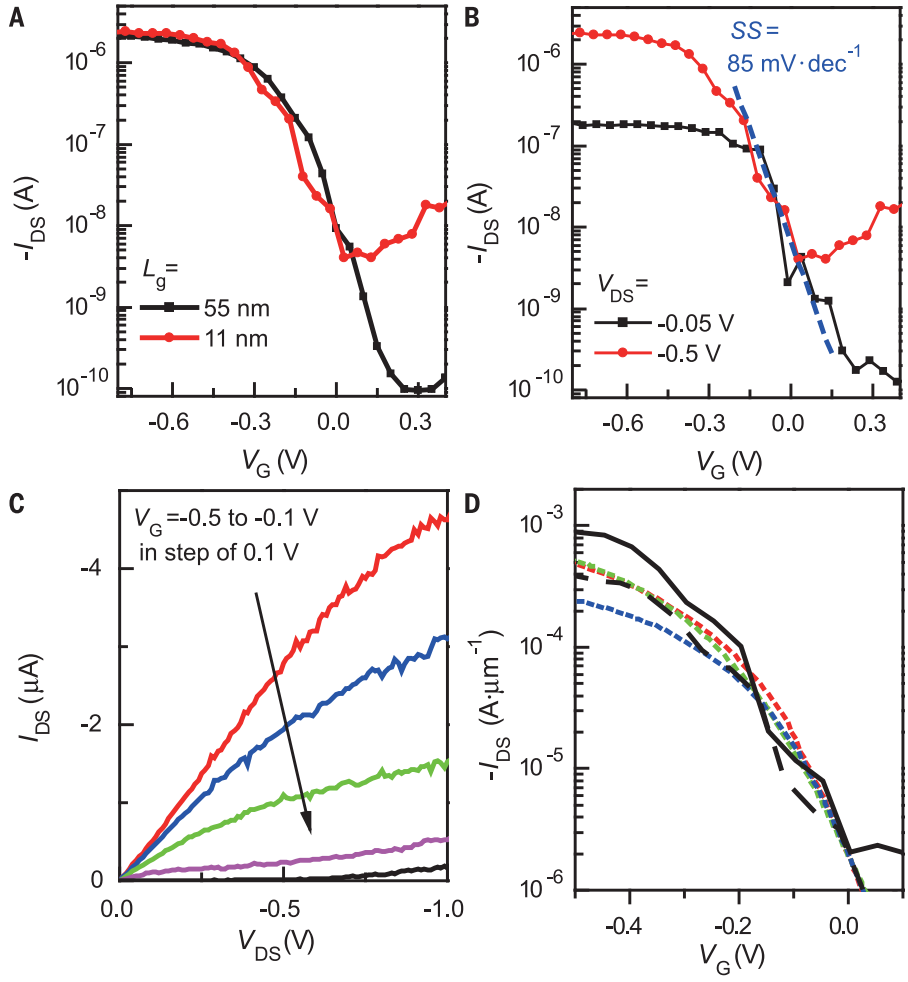

Fig. 2. Electrical properties of the individual s-CNT transistor scaled to the 40-nm device footprint. (**A**) Transfer characteristics of the two s-CNT transistors with identical L_c of 10 nm and W_{sp} of 5 nm but L_g reduced from 55 nm (black square) to 11 nm (red circle), showing an increase in minimal current but almost identical *SS*. V_{DS} = −0.5 V; V_G defined as how much gate bias is applied above the voltage to obtain I_{off} = −4 nA. (**B**) Subthreshold curves for the 40-nm-footprint individual s-CNT transistor under low (−0.05 V, black square) and high (−0.5 V, red circle) V_{DS} bias. (**C**) Output characteristics of the device as in (B) with V_G varied from −0.5 V to −0.1 V in step of 0.1 V from top to bottom. (**D**) Benchmarking pitch-normalized 40-nm-footprint s-CNT transistor scaled for the 3-nm technology node and beyond biased under −0.5 V V_{DS} (black solid line for 2-nm and dashed line for 5-nm s-CNT pitch, respectively) against the 10-nm-node Si FinFET biased under −0.75 V V_{DS} (red dotted line) (*20*), the 14-nm-node FDSOI FET biased under −0.8 V V_{DS} (green dotted line) (*22*), and the 10-nm-node Si nanowire FET with gate-all-around configuration biased under −1 V V_{DS} (blue dotted line) (*21*).

both the gate oxide thickness (t_{ox}) and the width of the spacer (W_{sp}); its profile determined L_g, which was 11 nm (Fig. 1B). Devices with single components smaller than the ~10-nm size for each critical device component in our transistor have been fabricated on various materials (*9*, *14*–*17*), but their overall device footprints were all much larger than 40 nm, with substantial surrounding hardware, including extended metal contacts or even scanning tunneling microscopy probes. The top-view scanning electron microscopy (SEM) images of a group of such devices (Fig. 1C) illustrate the straightness of the s-CNT and its good alignment with the device structures, assuring precisely defined device geometries. Each group contains two sets of transistors with different L_g fabricated on the same s-CNT to avoid variations caused by nanotube diameters. The L_g for devices in one set was made deliberately large [55 rather than 11 nm, fig. S3 (*13*)] to check the scaling with L_g.

Subthreshold curves of two devices with different L_g are compared in Fig. 2A, with the transfer (Fig. 2B) and full current-voltage (*I*-*V*) characteristics (Fig. 2C) of the smaller 40-nm-footprint s-CNT transistor (L_g = 11 nm), all with a constant back bias of −6 V to electrostatically dope the spacer regions near the contacts [see supplementary text 2 and fig. S4 (*13*)]. These two devices exhibit consistently low subthreshold swing (*SS*) at 85 mV dec^{-1}, indicating that they are still immune to short-channel effects

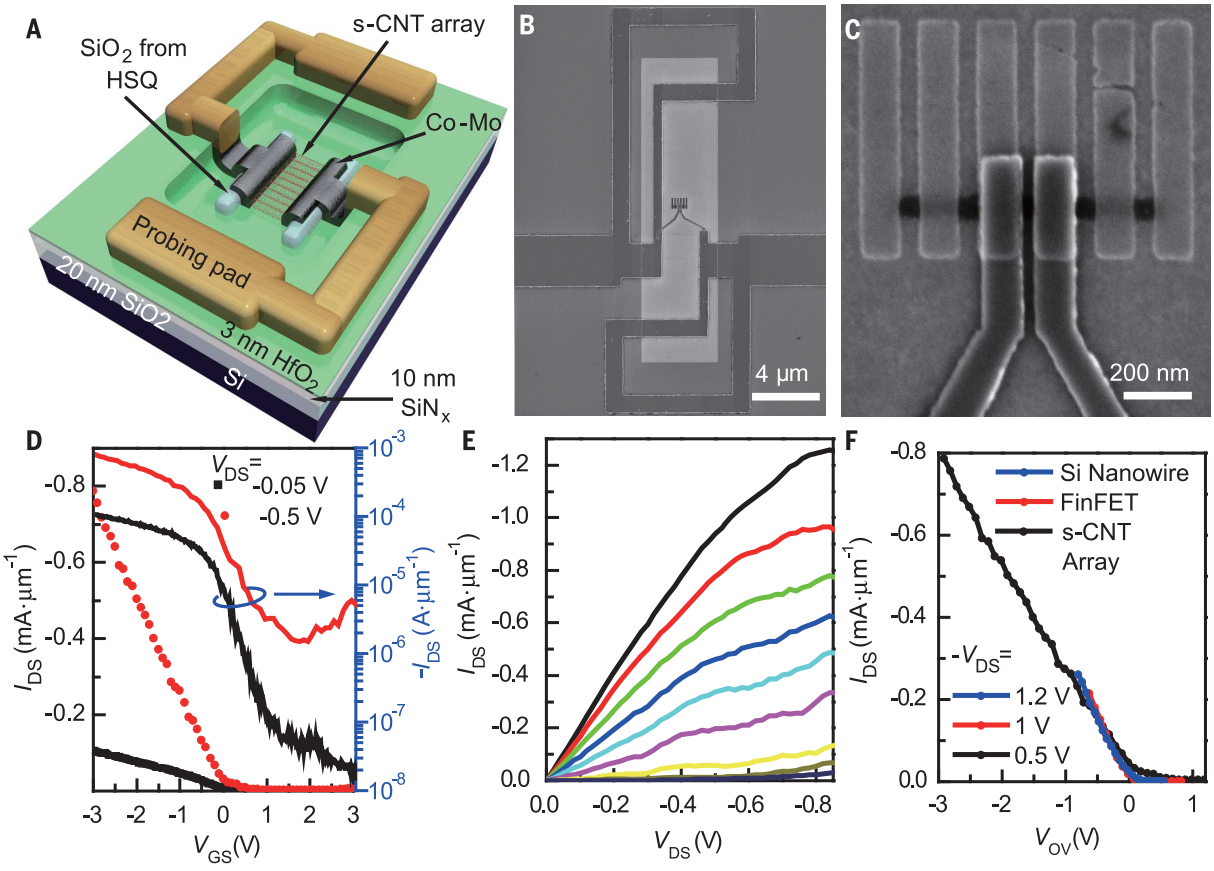

Fig. 3. Schematic, electron micrographs, and electrical properties of the scaled high-performance transistors built on s-CNT arrays. (**A**) Schematic showing a s-CNT-array transistor scaled to a 40-nm-device footprint with the channel sitting on a 3-nm HfO_2/Si back gate and the probing pads on 10-nm SiN_x/20 nm SiO_2 field oxide. (**B** and **C**) SEM micrographs under low magnification (B) to show the field oxide window and high magnification (C) to highlight the active device region. (**D**) Transfer characteristics for a scaled s-CNT-array device plotted in both linear (symbols, left axis) and logarithmic (lines, right axis) scales with applied V_{DS} of −0.05 V (black) and −0.5 V (red). (**E**) Output characteristics of the same device as in (D), measured with descending V_{GS} from −3 V at a step of 0.5 V to 1 V. (**F**) Benchmarking the scaled s-CNT-array transistor (black) with Si FinFET with 4-nm fin width (red) (*30*) and Si nanowire FET with 5-nm nanowire diameter (blue, I_{DS} normalized by 20-nm pitch) (*31*).

with 11 nm L_g and end-bonded contacts. They also exhibited almost identical on-current value I_{on} of ~2 μA, reflecting the quasiballistic transport where the device conductance is independent of L_g, at a low V_{DD} (the shared drive voltage $V_G = V_{DS}$) of 0.5 V with an off-state leakage current I_{off} of only 4 nA. Under this bias condition (V_{DD} = 0.5 V and I_{off} = 4 nA), compared with those s-CNT devices built with the same 10 nm L_g but more than 10 times longer L_c (*8*, *9*), I_{on} was only 1.5 to 2 times lower because the end-bonded contacts keep the contact resistance low even with such aggressively scaled contact size. Compared with the previous smallest footprint nanotube transistor with 20 nm L_g and 20 nm L_c built with palladium side-bonded contacts (*10*), I_{on} was identical despite the device's critical dimensions being a factor of 2 smaller.

Although some preliminary reports of 7-nm-node Si and SiGe FETs with device footprints of ~60 nm have appeared, their exact performance was not provided (*18*, *19*). Thus, we compare our devices to Si fins (*20*) and nanowires (*21*) scaled to the 10-nm node and fully depleted silicon-on-insulator (FDSOI) (*22*) scaled to the 14-nm node, the smallest FETs to date with quantitative reported *I-V* characteristics. With a roughly halved device size compared with 10-nm-node Si FETs, the nanotube FET still operates with similarly sharp *SS* (Fig. 2D). We normalized the current density by the nanotube pitch that we can currently achieve using full-coverage aligned arrays assembled by the Langmuir-Schaefer method (500 nanotubes per micrometer) (*23*) or projected to be attainable by the selective placement based on surface-chemistry-assisted self-assembly (200 nanotubes per micrometer) (*6*). Although the nanotube device was driven at a lower V_{DS} of 0.5 V, the normalized I_{on} values are near that (using placed tube arrays) or nearly twice as high (using full-coverage arrays) as that of the advanced Si devices under the same 0.5 V gate bias and 2 μA μm^{-1} I_{off}, showing the promise of extremely scaled nanotube transistors to simultaneously deliver higher device packing density and better performance.

Performance projections based on single s-CNT transistors as in Fig. 2D are valuable indicators, but the performance of devices built on actual arrays may be affected by nanotube variations (*24*) and other factors including Coulomb interactions and screening among closely packed neighboring s-CNTs (*25*, *26*). We fabricated and measured actual s-CNT–array devices in which high-purity s-CNTs were assembled from solution into well-ordered full-coverage aligned arrays as shown in fig. S5 (*13*) with the Langmuir-Schaefer method (*23*). These arrays were then integrated into bottom-gated transistors with Co-Mo end-bonded contacts and an overall device footprint still limited to 40 nm (Fig. 3A). For array devices, the inclusion of 50 atomic % Mo was especially critical for the formation of end contacts. Neither cobalt nor nickel forms a stable carbide phase or has enough carbon solubility to fully consume the underlying nanotube arrays (*12*). The gate dielectric was 3-nm-thick HfO_2, while probing pads sit on much thicker field oxide (Fig. 3B). The bottom-gate structure was used here to avoid the additional functionalization of nanotubes required for growing high-quality oxides by ALD on a fully covered carbon surface (*27*). The width of each device was patterned to be 50 nm (Fig. 3C). The detailed fabrication process is described in the supplementary materials and in fig. S6 (*13*).

The combination of better electrostatic control (from the thin gate dielectric), extremely high s-CNT density, and end-bonded contacts that allow low-resistance access to each s-CNT inside the array led to the unprecedented high performance for nanotube-array transistors even with extremely constrained device dimensions (Fig. 3, D and E). We obtained I_{on} up to 0.8 mA μm^{-1} with 0.5 V V_{DS} and an I_{on}/I_{off} ratio of ~10^3. The device transconductance (g_m) and the on-state conductance (G_{on}) were as high as 0.32 and 2.05 mS μm^{-1}, respectively, with a good yield of above 30% devices showing an on/off ratio above 100, reflecting the high purity of the s-CNT source [fig. S7 (*13*)]. Both values are more than three times as high as those of previous best-nanotube-array FETs [see supplementary text 3 and fig. S8 (*13*)] (*23*, *28*, *29*). The performance can be further benchmarked with that of advanced silicon transistors, including Si fin (*30*) and nanowire (*31*) devices with the smallest fin width and the nanowire diameter to ensure that their Si body is thin enough (two times further reduction from the 10-nm node) for the 5-nm technology node. As shown in Fig. 3F, the nanotube-array transistor delivers similar on-state current density under at least two times lower V_{DS} for the same gate overdrive (V_{OV}, defined as how much gate-source voltage (V_{GS}) was applied above the V_T, which was extracted using the standard linear-extrapolation method).

The one serious shortcoming of our array devices is the large *SS*, ~500 mV dec^{-1}, which leads to very poor subthreshold performance (I_{off} = 20 μA μm^{-1} at V_{OV} = −0.2 V) compared with current silicon transistors. Fortunately, the cause is already well understood (*24*). There is substantial variation of threshold voltage (V_T) and *SS* [measured standard deviation up to 530 mV for V_T and 260 mV dec^{-1} for *SS* as shown in fig. S9 (*13*)] commonly observed for individual s-CNTs, which arises primarily from the randomly distributed fixed charges in the gate oxide (*24*). For an array device, a few s-CNTs with large *SS* and low V_T could dominate the overall apparent device *SS* and limit it to a high value. Moreover, because of the V_T variation, the individual nanotubes do not all switch at once; they can only be turned on sequentially with the increase of V_{GS}, and this process further broadens the transition between device on-state and off-state. These two effects limit the *SS* for current array devices to ~400 mV dec^{-1}, according to our projection [see supplementary text 4 and fig. S9 (*13*)].

These fixed charges are caused by damage to the oxide by plasma and x-rays generated during the processing. We anticipate that their density can be dramatically reduced by further engineering optimizations, including better passivation schemes (*32*) and a replacement gate stack scheme to form the pristine gate dielectric in the last step of device fabrication (*33*). Previous simulations indicated that if this fixed charge density is reduced to a level comparable to that of a typical high-κ oxide/Si interface, the V_T uniformity of individual s-CNTS will be improved to a level suitable for very-large-scale integration (*24*). Although many other engineering challenges remain, such as achieving comparably scaled and performed n-channel nanotube transistors through previously developed physicochemical doping (*12*) or other approaches and forming arrays of such extremely scaled nanotube devices with 30-nm contacted gate pitch and shared source/drain contacts, it has been demonstrated that s-CNT transistors can actually be fabricated with dimensions and performance that make them a compelling candidate for a successor to conventional Si FETs.

REFERENCES AND NOTES

1. International Technology Roadmap for Semiconductors 2.0 2015 Edition (www.itrs2.net/itrs-reports.html) (2015). ITRS predicts that the device gate length should shrink to 10 nm, and the contact critical dimension to 11 nm, with 4-nm spacer width, to limit the overall device footprint to 40 nm in the 3-nm technology node and beyond.
2. C.-H. Jan *et al.*, *VLSI Symp. Tech. Digest*, T12–T13 (2015).
3. A. D. Franklin, *Science* **349**, aab2750 (2015).
4. A. D. Franklin *et al.*, *Nano Lett.* **13**, 2490–2495 (2013).
5. S.-J. Han *et al.*, *IEDM Tech. Digest*, 19.8.1–19.8.4 (2013).
6. H. Park *et al.*, *Nat. Nanotechnol.* **7**, 787–791 (2012).
7. M. M. Shulaker *et al.*, *Nature* **501**, 526–530 (2013).
8. A. D. Franklin *et al.*, *Nano Lett.* **12**, 758–762 (2012).
9. C. Qiu *et al.*, *Science* **355**, 271–276 (2017).
10. A. D. Franklin, Z. Chen, *Nat. Nanotechnol.* **5**, 858–862 (2010).
11. Q. Cao *et al.*, *Science* **350**, 68–72 (2015).
12. J. Tang, Q. Cao, D. B. Farmer, G. Tulevski, S.-J. Han, *IEDM Tech. Digest*, 5.1.1–5.1.4 (2016).
13. Supplementary materials are available online.
14. S. B. Desai *et al.*, *Science* **354**, 99–102 (2016).
15. M. Fuechsle *et al.*, *Nat. Nanotechnol.* **7**, 242–246 (2012).
16. J. Martínez-Blanco *et al.*, *Nat. Phys.* **11**, 640–644 (2015).
17. H. Lee *et al.*, *VLSI Symp. Tech. Digest*, 58–59 (2006).
18. R. Xie *et al.*, *IEDM Tech. Digest*, 2.7.1–2.7.4 (2016).
19. S. Y. Wu *et al.*, *IEDM Tech. Digest*, 2.6.1–2.6.4 (2016).
20. K. I. Seo *et al.*, *VLSI Symp. Tech. Digest*, 1–2 (2014).
21. S. Bangsaruntip *et al.*, *IEDM Tech. Digest*, 20.2.1–20.2.4 (2013).
22. O. Weber *et al.*, *VLSI Symp. Tech. Digest*, T168–T169 (2015).
23. Q. Cao *et al.*, *Nat. Nanotechnol.* **8**, 180–186 (2013).
24. Q. Cao, J. Tersoff, S. J. Han, A. V. Penumatcha, *Phys. Rev. Appl.* **4**, 024022 (2015).
25. F. Léonard, *Nanotechnology* **17**, 2381–2385 (2006).
26. J. Deng, H. S. P. Wong, *IEEE Trans. Electron Dev.* **54**, 3186–3194 (2007).
27. D. B. Farmer, R. G. Gordon, *Nano Lett.* **6**, 699–703 (2006).
28. M. M. Shulaker *et al.*, *IEDM Tech. Digest*, 33.6.1–33.6.4 (2014).
29. G. J. Brady *et al.*, *Sci. Adv.* **2**, e1601240 (2016).
30. J. B. Chang *et al.*, *VLSI Symp. Tech. Digest*, 12–13 (2011).
31. N. Singh *et al.*, *IEEE Electron Device Lett.* **27**, 383–386 (2006).
32. A. D. Franklin *et al.*, *ACS Nano* **6**, 1109–1115 (2012).
33. T. Y. Hoffmann, *Solid State Technol.* **53**, 20 (2010).

ACKNOWLEDGMENTS

We thank G. S. Tulevski, J. Bucchignano, S. J. Dawes, J. J. Yurkas, B. A. Ek, and A. Fustochenko for help with the processing. Q.C. conceived and designed the experiments. Q.C., S.J.H., and D.F. performed the experiments. Q.C. and J.T. performed device simulations. Y.Z. performed TEM analysis. Q.C. wrote the manuscript. All authors discussed the results (all of which are reported in the main text and supplementary materials) and commented on the manuscript.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/356/6345/1369/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S9
References (*34*–*52*)

20 March 2017; accepted 1 June 2017
10.1126/science.aan2476

Carbon nanotube transistors scaled to a 40-nanometer footprint

Qing Cao, Jerry Tersoff, Damon B. Farmer, Yu Zhu and Shu-Jen Han

Science **356** (6345), 1369-1372.
DOI: 10.1126/science.aan2476

Carbon nanotubes on the roadmap

The formal challenge for high-performance transistors is to fit within ever smaller devices. They need to shrink from a lateral dimension of about 100 to 40 nanometers. Cao *et al.* fabricated tiny devices by using a single semiconducting carbon nanotubes, as well as arrays of these nanotubes. High performance (a high saturation on-state current >1.2 milliamperes per micrometer and a conductance >2 millisiemens per micrometer) was delivered by making end-bonded contacts to the nanotubes with cobalt-molybdenum alloys.

Science, this issue p. 1369