

Replacing Silicon with Carbon Nanotubes

Why it's Still Worth Considering By Dr. Aaron Franklin

Since the inception of silicon (Si) in the 1960s as the principal material for fabricating integrated circuits, there has been interest in replacing it with a more ideal semiconductor. Dominating the various arguments for replacing Si is the imminent unfeasibility of further scaling (i.e., miniaturizing, shrinking) silicon metal-oxide-semiconductor field-effect transistors (MOSFETs)—the fundamental device of integrated circuits. Transistor scaling is driven by the desire to increase the density of transistors on a chip, thereby amplifying computing performance. Single-walled carbon nanotubes (CNTs)—consisting of a single atomic layer of hexagonally-arranged carbon atoms rolled into seamless cylinders of 1-2 nanometer diameter—have been among the foremost options for a Si replacement. This article briefly highlights the motivation for current projects at IBM that continue to consider CNTs for a future low-voltage, high-performance computing technology.

Death of Moore's Law for Si

In a 1965 research paper, one of Intel's eventual founders, Gordon Moore, made an observation and projection

regarding the rate of transistor scaling—his simple prediction actually became an industry-driving edict, known as “Moore's Law.” Briefly, the law states that the number of transistors on a chip will double approximately every two years (the timeframe has changed a few times between one and two years), as shown in the Figure 1 plot. While continued innovation has kept the Si MOSFET in stride with Moore's prediction, it is becoming more widely accepted that the end is close at hand—even within the next few years. What's most interesting is that the end of Si MOSFET scaling will not necessarily be caused by the inability to fabricate smaller transistors; rather, it is that the scaled transistors operate so poorly and at too high of power. As John Markoff in a recent New York Times article noted,

“The problem is not that they cannot squeeze more transistors onto the chips—they surely can—but instead, like a city that cannot provide electricity for its entire streetlight system, that all those transistors could require too much power to run economically.” [1]

In fact, it is safe to say that Moore's Law, as it was seen for the first forty years, has been dead for more than a decade now. As seen in Figure 1, deviation from the law has become common practice to meet specific application needs (high or low power/performance). Additionally, the push for faster and more efficient computing is no longer primarily tied to transistor scaling, but is becoming more and more stressed at the architectural level (e.g., multicore processors).

Why Carbon Nanotubes?

The struggle to keep Si in stride with projected transistor scaling is increasingly evident. The most recent major transition to keep Si alive [2] is a shift from planar devices to 3D fin (or trigate) structures for the 2012 technology. However, even if the Si fin is followed by the Si nanowire, the scaling limits for both gate length (around 10 nm) and supply voltage (around 0.8 V)—both key metrics to reduce for future technology nodes—have little hope of being overcome. It is true that doomsayers of Si have been around for years and innovation has continually managed to prove them wrong; but

with each innovation, the fundamental shortcomings of Si have become more apparent and increasingly more difficult to resolve. Consider the plot of power dissipation in the inset of Figure 1. We've reached the point where the static power (when a transistor is in the off-state) has become more costly than dynamic (on-state) power; this phenomenon is known as a leaky transistor and is largely a result of losing control over the current in a device at extremely small dimensions. For CNTs, where the channel body thickness is only approximately one nanometer (10 atoms), the transistor gate can control the current in the channel much more effectively, even when dimensions are aggressively downsized. This benefit can make for a more promising future trend in dynamic and static power dissipation, potentially lowering both curves seen in the Figure 1 inset.

One of the fastest growing application areas is low-power computing, wherein the supply voltage matters more than the transistor count, even at the cost of performance. Nanotube channels have shown the ability to operate better than any other demonstrated material at low voltages, making them a prime candidate for increasing transistor count even for low-power applications. However, the potential application space for CNTs is not limited to low power. The unmatched current-carrying capacity of nanotubes makes them equally as attractive for high-performance computing applications. Furthermore, there is an application space that to date is uncharted by commercial integrated circuits: flexible and/or transparent electronics. With nearly perfect transparency to visible light, high mechanical flexibility, and complete substrate independence, nanotubes lend themselves to a

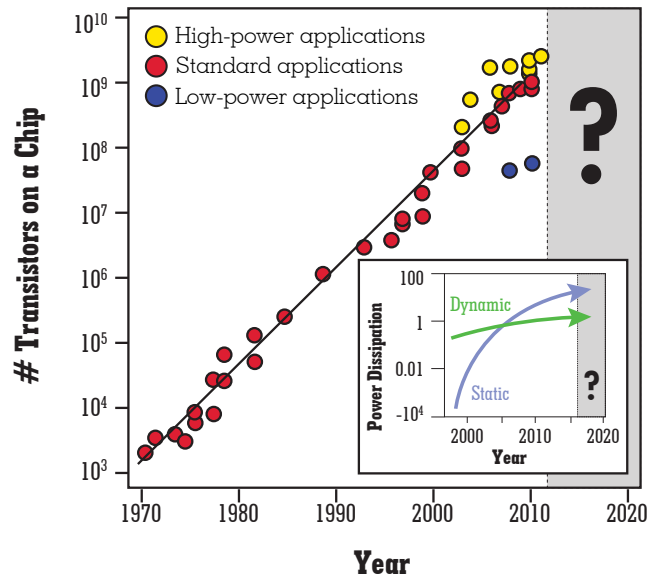


Figure 1: Number of transistors on a chip versus year of introduction, with transistor count doubling approximately every two years (known as Moore's Law). Each data point denotes a particular chip technology. Inset shows how the normalized total power dissipation scales with technology generations, yielding a serious power problem.

myriad of exotic applications that simply could not be realized with bulk semiconductors like Si.

IBM's Efforts Toward a CNT Technology

All of the promising attributes of a future CNT digital technology do not come without some substantial challenges. Foremost of the obstacles is the difficulty of placing nanotubes in precise locations with a consistent pitch. Attention to this problem has grown in the past few years, but much work remains to reach the goal of at least 100 CNTs per micron. Second is the need for improved material quality control. Because nanotubes can be semiconducting or metallic, obtaining material that is as close to 100 percent semiconducting as possible is a must (at IBM, we can reach 99.3 percent and counting). Finally, there is the need for improvements in the CNT transistor itself: optimized contacts, self-aligned gate structure, appropriate passivation layers, et cetera.

Despite the work that remains, it is amazing that CNT transistors have come so far in just over 13 years. Tremendous advancements have been made in both controlling the material and understanding/improving the device. Comparing the progress in CNT technology to that of the early bulk semiconductors is rather remarkable, as shown in Figure 2. Early development of transistor technology was a completely uncharted path, relying on the unpredictable advancements that occurred in the field. When Moore's Law came along, progress in the integrated circuit world began its march on a predetermined path. Since 1998, this path has been devised by a group called the International Technology Roadmap for Semiconductors (ITRS), which maps out the needed deliverables (e.g., device dimensions, performance) for technologies that are up to 15 years out [3]. Therefore, CNT transistors can't just be better than the current Si MOSFETs; rather, they must work

significantly better than any achievable Si device in order to justify the overhead cost of transitioning to a new platform. This is what we are working to determine right now at IBM.

At the IBM T. J. Watson Research Center, we are uniquely staffed and equipped to take on the remaining hurdles to a CNT technology. In-house, we synthesize, purify, and isolate large quantities of semiconducting CNTs. Strategies for improving precision in placement of the nanotubes onto 200 mm wafers are under intense study. We also continue to improve our understanding of CNT transistor performance and scaling limits, including the optimization of a technologically viable device structure. Does all of this mean we guarantee CNTs on the integrated circuit roadmap in the next few years? No. But what it does guarantee is that we will soon have a definitive answer regarding whether or not a CNT transistor technology is possible and practical enough to drive the next revolution in digital computing.

References

[1] http://www.nytimes.com/2011/08/01/science/01chips.html?pagewanted=1&_r=2&ref=science

[2] <http://newsroom.intel.com/docs/DOC-2032>

[3] <http://www.itrs.net/> ■

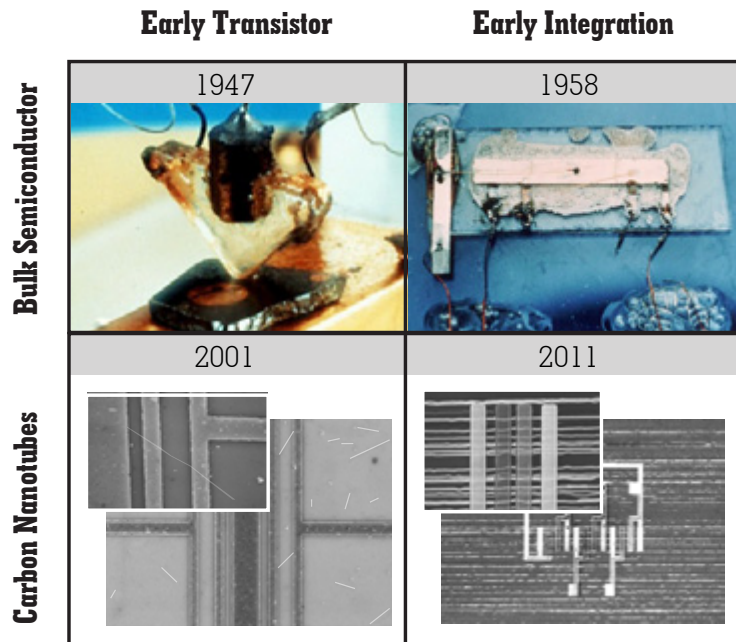


Figure 2: Early development of bulk semiconductor transistor technology compared to development in carbon nanotube technology. Pictures of the first demonstrated transistor (1947) and one of the first demonstrated integrated circuits (1958) are from <http://www.computerhistory.org>. Scanning electron microscope images of CNT devices are from work at IBM (the bright lines are CNTs), showing how far integration of nanotubes has come from random dispersion (2001) to parallel alignment (2011). The bottom right images are of fully integrated ring oscillators with multiple nanotube channels in each transistor.

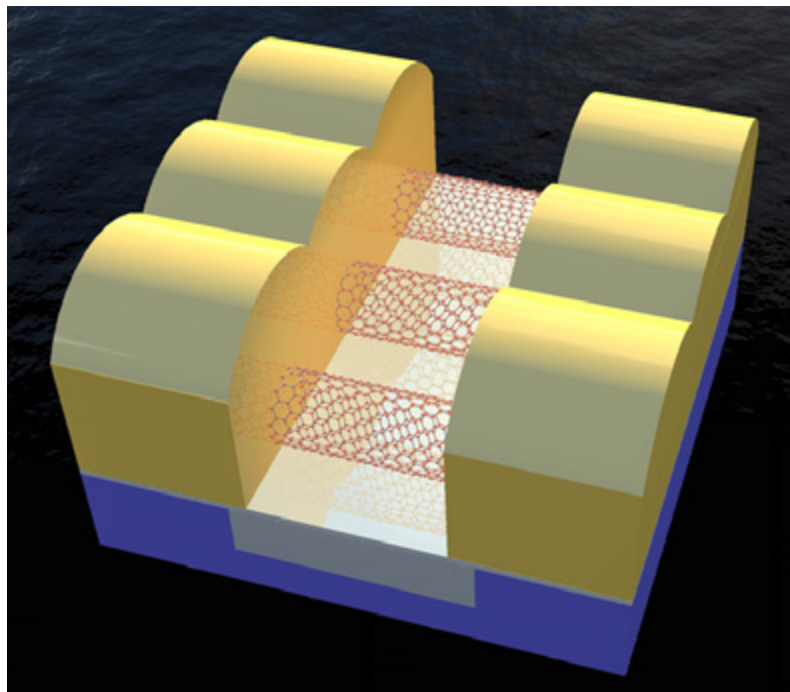


Figure 3: Schematic image of carbon nanotube transistor with multiple channels.