INTRODUCTION

Given the history of the semiconductor industry, most of these issues can probably be solved with current processes. However, there are two significant exceptions. Physical size limitations and astounding costs may require a shift in the fundamental way integrated circuits are fabricated. Many researchers believe this shift will be to nanoelectronics.

With a mix of chemistry, physics, biology and engineering, nano-electronics may provide a solution to increasing fabrication costs, and may allow integrated circuits to be scaled beyond the limits of the modern transistor. The largest change in a shift to nano-electronics is the method of fabrication. Individual wires, diodes, field effect transistors (FETs), and switches can be created abundantly and cheaply in a test tube. All of these devices are only a few nanometers in size, and may reach a level of integration not possible with conventional ICs. It is estimated that nanoelectronics will be able to integrate $10^{12}$ devices per cm$^2$, while the ITRS estimate that at the end of the roadmap in 2018 manufacturers will only be able to achieve 1010 MOSFET transistors per cm$^2$. This level of integration will be difficult to achieve due to the components’ miniscule dimensions. It might be impossible to individually pattern the small components of the nano-electronics in the ways that current fabrication processes allow. While current ICs can have almost any arbitrary pattern, nano-electronics will likely have a regular structure generated by a stochastic self-assembly process. Unlike deterministic self-assembly, stochastic self-assembly means that chips will be fabricated with methods that allow components to guide each other in constructing a structure with little or no outside intervention. This is often referred to as a “bottom up” method, because the individual parts are built and then assembled into an architecture, and the use of the architecture is based on available resources. This is in contrast to a “top down” method used in current IC fabrication, where designs are conceived at a high level and the necessary components are put together to implement the design. The lack of outside intervention means that fabrication is more prone to defects and no single part can be absolutely relied on to be functional. In current lithography based electronics, the most popular model for handling defects is to reject any chip with even a single defect. This model will no longer work with nano-electronics because their defect densities will mean that no chip will be totally defect free. This suggests that nano-electronics will likely need to be reconfigurable like an FPGA in order to function in spite of defects. In this paper, we consider the major research efforts for nano-electronics by surveying proposed technologies for replacing the transistor, possible chip architectures, techniques for handling defects, and software implications. We focus on the higher-level electronic aspect of these topics, though we provide...
CNT electrical devices:

Currently, the most promising use of semiconducting CNTs is as a transistor component. As can be seen in Figure 1, carbon nanotube field effect transistors (CNTFET) appear very similar to MOSFETs, with the silicon channel replaced with a CNT. Most of the CNT transistors have been fabricated with SWCNTs because their band gap energy is in the range of a semiconductor. One group, however, found that MWCNTs could be used if the nanotubes were collapsed or crushed. This is probably impractical for large-scale systems, since each nanotube would have to be individually collapsed or selected amongst many “normal” nanotubes. Two varieties of CNT transistors have been fabricated (Appenzeller, J. et al., 2002). Figure 1a shows an illustration of a CNT transistor with a back gate (gate placed under the channel instead of over it), which uses the silicon substrate to control the conduction through the CNT. The use of a back gate is easier to fabricate, but it has the disadvantage of not being able to control the individual transistor because the substrate is shared between all transistors. This configuration is good for research, but is probably not a realistic candidate for commercialization. The other variety uses a gate that is over the top of the CNT, as in Figure 1b. These so-called second-generation CNT transistors have two advantages over their counterparts with a back gate. The most obvious advantage is the ability to individually control the FETs because the gates are isolated. The gate on top also allows for a thinner gate oxide, which means that the controlling voltage can be lower. Also, CNTs are intrinsically p-type, but they can be altered to behave as an n-type semiconductor [Nosh05], however, exposing an n-type CNT to oxygen will cause it to revert back to its native p-type. Covering the CNT with the gate is a good means to isolate it from oxygen. Individual gating and the formation of both p-type and n-type allows for CNT transistors to be arranged in complementary pairs much like current CMOS. Unfortunately, it is much more difficult to fabricate these transistors with a top gate. (Bachtold, A. et al., 2001)

CNT based transistors have promising enough characteristics to prompt companies such as Intel, NEC and IBM to investigate them as replacements for modern transistors. The first advantage is the small size of the CNT. The small diameter of the CNT means that all parts of the channel are close to the gate, and they are easier to control. Another advantage of using CNTs is that they exhibit ballistic transport of electrons because of the tube structure. Since all of the atoms in the tube are bonded to the same number of neighbors, there is no electron backscattering. This is in contrast to a wire made of a crystal, which has irregular bonds at the surface. Ballistic electron transport means that transistors with CNTs will exhibit higher on currents that will not be affected by the length of the transistor channel (Bahar, R.I., et al., 2004). For MOSFETs, the current decreases as the channellength (distance between the source and drain) increases. An unsolved problem with the use of CNTs for the transistor channel is increasing the width of the channel. For MOSFETs increasing the width of the channel increases the current drive capabilities of the transistor, which is absolutely critical for circuit design. With CNT transistors, the only way to achieve this would be to “lay” CNTs side by side, since the tube dimension is set. Unfortunately, there is currently no technique for performing this. Although CNT transistors and the MOSFETs discussed in section 1.1 behave alike and appear very similar in structure, the operational physics are very different. The CNT is not in contact with the bulk to transfer carriers, as is done with MOSFETs. The transistor behavior arises from Schottky barriers at the source/CNT interface and its interaction with applied electric fields. Schottky barriers are formed when a metal and a semiconductor are joined together, and there is an energy difference between the Fermi level (EF) of the metal and the energy level of the carrier (holes or electrons) of the semiconductor. The Fermi level is the top energy state possible for an electron in the metal at 0 Kelvin. When the Fermi level of the metal is between the conduction (E_c) and valence (E_v) band of the semiconductor, carriers have to acquire energy to move between the source and the semiconductor. In order to clarify the process of how the transistor is turned on and off, an example of a p-type CNTFET is given in Figure 2. Figure 2a shows the band energies of the CNTFET without any voltage stimulation. The Fermi levels
of the source and drain are different because the positive voltage on the source lowers the energy level and raises the Fermi level of the drain. When there is no bias on the gate, the Fermi level of the source is higher than the energy level of the holes in the valence band of the CNT. This barrier means that very few electrons can move from the holes in the CNT to the source even though an electric field exists between the source and drain.

(even though holes are carriers in p-type transistors, it is electrons moving between the holes in the valence band that actually create the current). When a negative bias is placed on the gate in Figure 6b, the valence and conduction bands are raised. Except for a small portion near the source/CNT interface, the valence band is above the Fermi level of the source. This means that the Schottky barrier is very low and electrons easily tunnel from the CNT valence band to the source because they are in a higher energy state for most of the valence band. The process is the same for an n-type transistor, except that now the electrons move through the conduction band, the Schottky barrier would be on the drain side of the diagrams in Figure 2, and a positive bias on the gate would lower the conduction and valence bands in Figure 2b. (Bhaduri, D., S.K. Shukla, 2004)

Another promising application that takes advantage of CNTs strength properties instead of their electrical properties is as non-volatile memory devices. The first proposal was an array of SWCNTs with contacts at one end of each CNT (see Figure 3). One layer of CNTs sits on the substrate while the other layer is suspended over the first layer by a spacer. To write to the memory, opposite charges are placed on two orthogonal CNTs. The two contacting CNTs now have a non-infinite resistance between each other, and are considered on or ‘1’. Locations where the CNTs have not been bent, and thus there is no connection between the perpendicular CNTs, are a ‘0’. To read a cell, current is sent down one CNT; if current is detected on the output of the orthogonal CNT, the two CNTs are making contact. A like charge can be placed on two contacting CNTs to separate them and erase a ‘1’. Mechanical forces will keep the two CNTs separated when the like charges are removed. The fact that the CNTs stay in their configuration without electrical charge due to van der Waals or mechanical forces makes this memory nonvolatile. (Brown, C.L. et al., 2000)
The RAM in Figure 7 requires two layers of CNTs, with a placement of the top layer over the spacers. This is a difficult task with CNTs, so the design was modified to only have one layer of CNTs, which are suspended over metal electrodes (see Figure 4). The metal electrodes are arranged in long troughs, and the CNTs are placed orthogonally over the troughs, eliminating the need for exact placement. To increase the robustness of the memory, each cell contains multiple CNTs connected to a contact. The read/write procedure is identical to the above architecture. (Brown, J.G., R.D. Blanton, 2004)

![Fig. 4: Cross-section view of a CNT memory cell with metal electrodes.](image)

**Conclusion:**

The invention of the transistor in 1947 is one of the most important inventions of the 20th century. Since its inception, the transistor has been reduced so that now modern devices are orders of magnitude smaller than their earliest counterparts. Unfortunately, the scaling down must eventually end. Increasing power, capital costs, and ultimately theoretical size limitations, are poised to halt the process of continually shrinking the transistor. Nano-electronics show promise as a technology to continue the miniaturization of ICs. However, whether nano-electronics will be a replacement for conventional ICs, or as a complimentary technology, is yet to be determined. What has already been shown is that components such as wires and molecular switches can be fabricated and integrated into architectures. It is also known that these devices will be prone to defects and that fault tolerance schemes will be an integral part of any architecture. Finally, the preliminary research indicates that while existing parts of the CAD tools will be useful for nano-electronics, there will need to be some additions and changes made.

The greatest progress has been made in the research of the components that may make up nano-electronics. Chemists have been able to fabricate molecules that have two states, such that the molecules can be switched “on” and “off”. Some of these molecules have shown the functionality of diodes or variable resistors. Chemists have also been able to fabricate silicon nanowires and carbon nanotubes. Both of these technologies can be used as wires or devices, and in some cases both. Nanoinprint lithography, probably the most promising wire fabrication technique, has been used to produce working memories on the nanometer scale. While all of these devices have been demonstrated, more lot of research is required to reliably produce these devices, and to create better devices. One of the big questions for the future nano-electronics is whether nano-scale devices can be reliably assembled into architectures. Some small-scale successes have been achieved, but the benefit of nano-electronics is the enormous integration levels they may be able to achieve. The most promising architectures to date are array based. This is because arrays have a regular structure which is easier to build with self-assembly. Arrays also make good use of the available devices (nanowires, carbon nanotubes, and molecular electronics), and they are easy to configure in the presence of defects. There are other more random architectures that would require even less stringent fabrication techniques, but there is some doubt about how they will scale to larger systems. Overall, it is difficult to evaluate architectures as the underlying components are not fully understood nor developed yet. One thing that seems clear is that nano-electronics will, at least for the first few generations, need the support of conventional lithography based electronics for things such as I/O, fault tolerance, and even simple signal restoration. Fault tolerance is another big problem for nano-electronics. It seems evident that the manufacturing techniques may never be able to produce defect free chips, so fault tolerance will be key to the success of nano-electronics. For manufacturing defects, detecting and configuring around the defects is the most economical technique, since nano-electronics will be configurable devices. The hard problems are detecting the defects among 10^12 devices in an economical manner, and how to best manage the large defect map. It also appears that transient faults will be a problem with nano-electronics due to their small size and low current levels. To handle transient faults, a hardware redundancy method such as multiplexing or NMR will have to be used to dynamically detect and repair faults. Unfortunately, these methods would require too much redundancy to handle the number of manufacturing defects expected. One aspect of nano-electronics that resembles current
technologies is their CAD flow. Much of the software for utilizing nano-electronics will resemble that of FPGAs. A nano-electronic CAD flow will still have technology mapping, placement and routing to produce configuration files, plus some additional steps. The additions will be a routine to detect the defects before placement, and some kind of backend step to handle the unique circumstances surrounding deployment. The big issue is how to deploy a circuit on a nano-electronic chip when each chip is unique. With current reliable devices, one design can be used to produce millions of chips. If nano-electronics are to become more than a niche computing tool, a deployment model must be developed that doesn’t burden the end user or cost the manufacturer excessive testing time. As can be seen, a substantial amount of research has been conducted on nano-electronics. Many working devices have been designed and fabricated, along with a number of small-scale memory chips, but there are some big hurdles to overcome. These hurdles include lowering defect levels to a point that reasonable redundancy levels can be used, integrating billions of devices, and developing software tools to complement the new technologies. However, the prospect of cheaply integrating $10^{12}$ devices per chip is a powerful incentive to overcome the challenges. With a little more than 10 years before the projected end of scaling for lithography based circuits, answers to these questions will hopefully come within the decade.

REFERENCES


