Discussion of Benefits and Concerns Regarding Application of Graphene Nanotransistors to Affect the Computing Energy Issue

Anthony Nguyen, Brian Aller, Dom Venturino

University of California, Los Angeles

1. ABSTRACT

Computer energy consumption is a real environmental issue that we must address in the upcoming decades of the 21st century. Energy use doubling combined with the increased proliferation of computing technology means that we must start using efficient machines, and now. Transistors are a fundamental component of computer systems, which are responsible for the computer's calculations, and the prime consumer of energy. Graphene Nanotransistors are a novel version of the traditional transistor circuit element, with monolayer carbon as an exciting new channel material. There are two problems associated with traditional transistors that our Graphene Nanotransistors improve upon: performance issues due to Switching Speed and lossy energy consumption due to Leakage Current. As demonstrated by IBM and UCLA, Graphene Nanotransistors have higher electron mobility than silicon, and the behavior of graphene bypasses some of the issues involved with miniaturizing traditional transistors. Algorithms developed at UCLA show that even when we evaluate the benefits of graphene conservatively, performance and energy savings are very significant. Studies about the environmental and health impacts of graphene show that the gains outweigh the costs, although more research is necessary in this area.

2. INTRODUCTION--What the Problem is

Computer energy consumption is a real threat, one that may one day join the threatening ranks of oil and water if we don't act, and act fast. Energy consumption from computers doubled from the year 2000 to the year 2006 (*1*). According to Greenpeace, "If the Internet was a country, it would rank fifth for the amount of electricity usage, just below Japan and above Russia" (*2*). And the energy demand needed to power the companies leading the Internet revolution is growing at a rate of 12 percent annually. Half of the 61 billion kWh of energy used by IT companies in the course of the EPA's 2007 report was put into infrastructure, such as cooling and power delivery (*3*). One of the giants in the computer driven area, Google, says it used 2.3 billion kilowatt-hours of electricity last year, about the same as what 207,000 U.S. homes would use in a year or the power consumption of about 41 Empire State Buildings, according to Edison Electric Institute, a utility industry trade group (*4*). The energy usage is so high, that these companies need to find power plants to their stations. Servers, which are under constant

CPU load and are processing data all day long, need more efficient CPUs. And to answer that calling, we are here to discuss a fundamental building block of the CPU, the transistor.

3. TRANSISTORS

We now move to a discussion of the basic elements and functioning of a standard transistor. This will illuminate our work on graphene transistors, and provide less experienced readers with a reasonable background.

3.1 – What Is a Transistor?

Transistors can be considered to be the "neurons" of the computer chip; it is in the conceptually-vast yet physically-tiny network of millions or billions of transistors that most of the real work of the chip occurs. "Computers are made up of 0's and 1's", they say, and it is in the transistors where all these 0's and 1's (called bits, and represented by high voltage to represent Binary State 1, and low voltage to represent Binary State 0) are propagated, and changed (called flipping a bit, from 0 to 1, or 1 to 0). When many transistors are placed together in a designed network, and operate in concert, this propagation and flipping of bits almost magically can become complex computation, high-level analysis, stunning graphics, user-interaction, and any and all of the features of computing systems we love so much.

3.2 – How Transistors Work – A Basic Model

Transistors, specifically the Field Effect Transistors (FETs) which are used in computers, can be abstracted to having a Source, a Drain, a Channel (shown red in *Figure 1*) connecting the Source and the Drain, and a Gate which controls the Channel. Because of a voltage difference between the Source and the Drain, current is "inclined" to flow between them. However this current is normally blocked because the Channel region between the Source and the Drain does not permit conduction, and thus the transistor is considered OFF (Binary State 0). However when a voltage is applied to the Gate, it in effect "opens" up the Channel by allowing conduction and current is allowed to flow (*3*), and thus the transistor is considered ON (Binary State 1). This current is either negative or positive depending on whether the transistor is n-type or p-type, but the difference between the two types bears little on our discussion. The change in the voltage of the Gate in effect "switches" the transistor from ON to OFF and vice-versa.

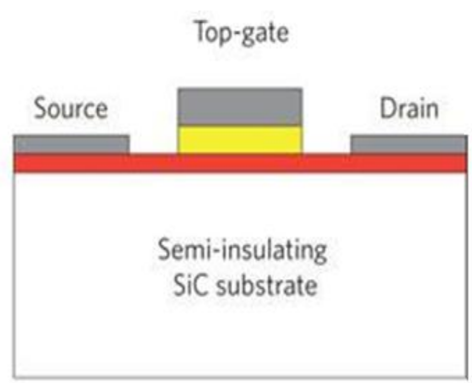

Figure 1: Basic Field Effect Transistor Model

3.3 – Optimal Transistor Characteristics

We have idealized a number of aspects of transistors in the previous section in order to present a clear model, and now explicate the real-world nature of those aspects that pertain to our nanotechnological application, starting with a discussion of Leakage Current.

In reality, when the transistor is OFF, the current between the Source and Drain is not actually 0, but can be any amount, provided it is noticeably lower than the ON current. Fortunately, using binary states does not require absolutes—it only require that 2 states can be distinguished. This current that is flowing during the OFF state, or when the transistor is non-active, is called Leakage Current, and it serves no functional purpose. Leakage Current is one of the prime factors contributing to energy consumption in a computing system. An optimal transistor has as low a Leakage Current as possible, but due to physical realities, this can be very hard to achieve.

Another consideration is Switching Speed. In reality, switching a transistor from ON to OFF is not an instantaneous process. Two main factors come into play in the Switching Speed. First, the electrical properties of the Channel region must be changed by the Gate voltage in order for switching to occur, and not all channel materials change at the same rate. Secondly, the electrons have to actually move from one side to the other across the channel in order for current to flow and for the transistor to register as ON, and electrons move different speeds in different channel materials. This electron speed is called Electron Mobility. An optimal transistor has as high a Switching Speed as possible, which often involves as high an Electron Mobility as possible (*4*). The speed of each transistor contributes to the overall clock cycle time of the circuit, which is the time the circuit must wait for a signal to pass beginning-to-end through the slowest path (Critical Path) in the circuit. This wait time is called the Critical Delay and it is during this wait that leakage is occurring, thus lowering the Critical Delay is both speeding up the circuit and decreasing leakage.

4. GRAPHENE NANOTRANSISTORS

Graphene Nanotransistors are perhaps the most exciting new development in the transistor industry (*5*). They offer many exciting possibilities, the first of which is significant speedup, another of which is decreased energy consumption.

4.1—What is Graphene

Graphene is a single-atom thick layer of carbon atoms in a lattice structure. It has extraordinary properties, most relevant to our discussion being unprecedented Electron Mobility orders of magnitude higher than that of Silicon, and the highest of any known material at room temperature (*6*). Graphene, being only a single atom thick, is also orders of magnitude thinner than standard channel materials, the benefit of which is explained in the next section.

4.2—Why Graphene Transistors Offer Potential Improvement

Graphene Nanotransistors can have much faster Switching Speeds than traditional transistors, as one can imagine by Graphene's incredible Mobility. In fact researchers at UCLA led by Xiangfeng Duan recently achieved 300 GHz Graphene Nanotransistor speed (*7*). For comparison, the newest MacBook Pro with Turbo Boost reaches 3.6GHz (*8*).

When normal transistors are scaled down to increase speed (smaller transistor = shorter channel, and shorter channel = shorter delay for current to cross channel), debilitating Short Channel Effects occur which result in Leakage Energy significantly increasing (*9*). However because Graphene is so incredibly thin, transistors with Graphene as the channel material manage to avoid short-channel effects (*5*) and thus have a significant reduction in Leakage Energy.

But an important caveat to this decreased Leakage Energy is that Graphene in its standard state doesn't have a Band Gap. A Band Gap is what allows a channel material to alternate between allowing conduction and preventing it. Without a Band Gap, a transistor would be always ON and have immense Leakage Energy. In exchange for sacrifices in speed, Band Gaps can be introduced to Graphene by making Graphene Nanoribbons (GNRs) (*10*). There is debate, however, as to whether or not Graphene Nanoribbon Field Effect Transistors (GNRFETs) will easily be able to both be significantly faster and at the same time have lower leakage energy than traditional transistors (*11*) (*12*).

5. ALGORITHMS AND APPLICATIONS

UCLA researchers Miodrag Potkonjak, Brian Aller, and Saro Megeurdichian, began research in 2011 on applications of Graphene Transistors in modern circuits (*13*). They developed a simulation and optimization system for the integration of GNRFETs into modern circuits based on algorithms they developed. They used very conservative values for Leakage characteristics of Graphene Nanotransistors in their modeling and simulations. This decision was because current research is divided on the matter and it is more compelling research if an innovation can be shown to work even in the worst case. Furthermore, the most aggressive estimates for Graphene Nanotransistors are so good that no research is required to prove such transistors would be improvements. Values used for speedup of a transistor due to Graphene ranged from 10x to 100x. Values for increased Leakage Energy of a transistor due to Graphene ranged from 10X to 10000X. One aspect of their research was also in the heterogeneous aspect of their design, which has been gaining popularity in computer hardware design, especially in Domain Specific

Computing (*14*), but had yet to be considered with Graphene in mind.

5.1—Basis

A computer chip may have a billion transistors on it, but there's no reason why they all need be identical or homogeneous. In fact, as researchers have been discovering more and more, there can be great advantages to strategically-engineered variations between different transistors (a type of heterogeneity). In any given circuit, transistors will have different workloads from each other, either by having different amounts of inputs, different amounts of outputs, or by what percentage of the time they are in-use. There is a strategy called "gate sizing" that involves changing the size of certain transistors to make them slower and consume less energy, or faster but consume more energy, depending on what type of workloads are on the transistors and the nature of their connections to the rest of the circuit (*15*). One algorithm for Graphene Nanotransistor-based heterogeneity uses gate sizing as its jump-off point, but involves heterogeneity between traditional Silicon transistors and ones made of Graphene rather than heterogeneity of transistor size.

5.2 – Algorithmic Summary

So how is it that substituting in Graphene transistors that consume More energy can result in a circuit that uses Less energy? The precise algorithms developed are proprietary by Aller et al. are proprietary, but the basic idea is as follows. If the entire circuit were to be speeded up magically, say to cutting the clock cycle in half, the leakage energy would be cut in half, since the each transistor would be spending half as much time leaking. But if the circuit were sped up by making every transistor high leakage high speed (conservatively estimated) Graphene Nanotransistors then high intrinsic leakage of each transistor would outweigh the savings from the clock cycle cut. However, if a just small number of Graphene Nanotransistors are algorithmically placed in ideal locations, such as in bottlenecks, then large speedup can be achieved and Leakage Energy still decrease. The difficult part is figuring out how much of the circuit to make Graphene and which among the millions or billions of transistors should be substituted.

6. RESULTS

Initial results from their algorithms, shown here for a 100 million transistor circuit in Figure 2, using the worst of the worst case where each Graphene Nanotransistor has 10x Speedup and 10000x Leakage compared to a normal transistor, have shown almost a 25% reduction in overall energy consumption coupled with an almost a 6 times increase in speed when Graphene is algorithmically placed. Since these results are using the conservative estimates, this shows that Graphene Nanotransistors can bring substantial improvements in energy consumption to computer circuits, even if they don't measure up to what they are well expected to.

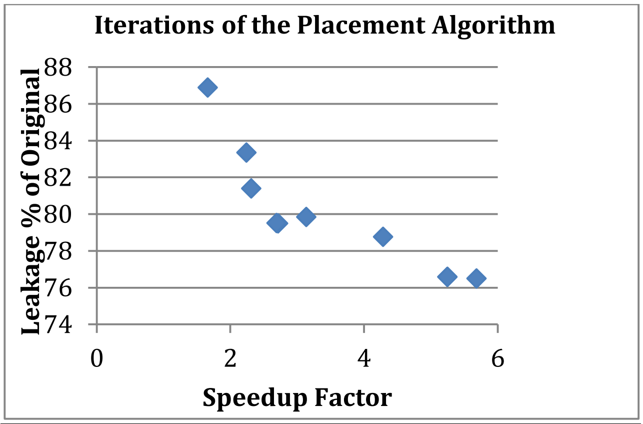

Figure 2: Results for 10x Speedup 10000x Leakage Graphene Substitution over 100,000,000 Transistor Network

6.1 --Environmental Projections

A 25% reduction in leakage energy would already save more than 15 billion kWh of electricity per year in servers data centers alone, and 7.5 million tons of CO2 kept out of the environment. And these predictions, once again, are based off very conservative estimates from currently available Graphene Nanotransistors. These predictions, and the research behind them, serve to show the truly mountainous effect Graphene Nanotransistors could have in the near and far future.

7. Environmental Impact of Graphene

While graphene nanotransistors offer great efficiency for large-scale computer processing, the effects of nanoscale graphene on the environment must be more fully understood before they are mass-produced. It is true that environmental regulations are not yet applied to nanomaterials, but from the standpoint of ethical science and engineering, the environmental impact of any technology should be studied and its safety should be maximized. Luckily, because of the way graphene is manufactured and produced in transistors, there is very little chance of human exposure to it, especially when compared to nanomaterials used in construction applications where demolition occurs, like titanium oxide. Ultimately then, it is the disposal and potential recycling of CPU towers that will act as the greatest chance of environmental exposure.

Because the field of nanotechnology is still fledgling—it has only been developed extensively since 1986 (SOURCE)—the environmental studies that have been conducted and published are typically done for specific nanomaterial-test subject combinations. Therefore, though there have been numerous tests done on graphene, the tests themselves are done on a single type of cell, and the typical testing method is *in vitro*. Graphene oxide (GO) is a nanomaterial that is often studied due to its range of applications, from its use for the production of graphene to its use as a strong paper-like material. Though graphene oxide may not be the principle type of graphene used in nanotransistors, its effects tested on eukaryotic cells are likely to be similar for a range of different compounds falling under the graphene category. Some of the cells with which graphene is tested are A549 cells—adenocarcinomic human alveolar-based epithelial cells; BEAS-2B cells—human lung cells; and Vero cells—kidney epithelial cells extracted from an African green monkey. It's likely these types of cells are chosen because they would be the cells affected if graphene were to be inhaled.

The studies to be analyzed are two done on graphene: one using graphene oxide, which is a just a sheet of graphene treated with oxidizing agents on its surface, done by the Institute of Nanochemistry; and Nanobiology at Shanghai University, and one comparing pristine graphene and carboxyl-functionalized graphene done by the Amrita VishwaVidyapeetham University in India.

7.1—Graphene Oxide Examination

Figures 3 and 4 show the results of the graphene oxide-A549 cell study. (*16*)

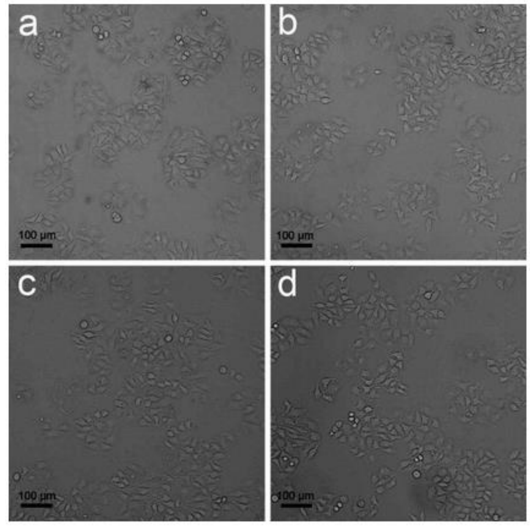

Figure 3: A549 cell morphology

Figure 3 above shows the cell morphology of the A549 cells treated with a) large particle GO samples, b) small particle GO samples, c) mixed particle GO samples, and d) the untreated control group. Cell structure is often a telling indicator of the status of the health of the cell, and since all four images show a very similar structural pattern, none of the GO introduced had a significant negative change on cell structure.

Figure 4 below shows cell viability of A549 after 24 hours when introduced with the increasing concentrations of the large-, small-, and mixed-particle-size GO.

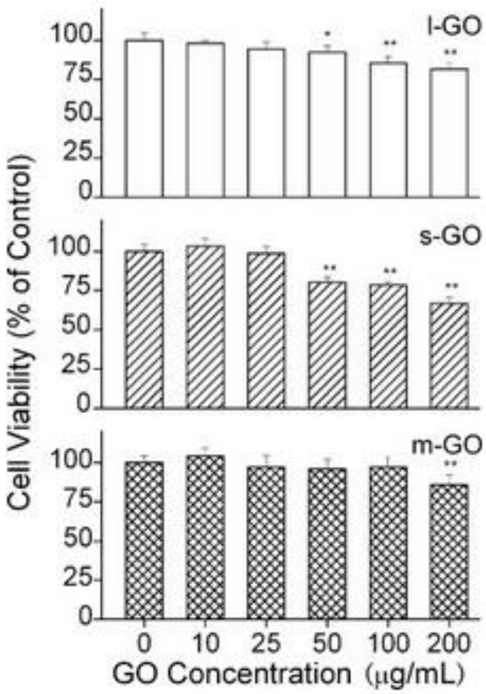

Figure 4: A549 cell viability vs. GO concentration

In all three cases the cell viability is dose dependent, meaning that there is a loss in cell viability in all three type samples around 50 micrograms per milliliter.

Because of the frequency of the production of reactive oxygen species—molecules that tend to cause destruction in living cells, typically abbreviated as ROS—for carbon-based nanomaterials, the ROS production caused by graphene oxide was measured for this test. Figure 5 below shows the results of the ROS concentration test. The units for measuring ROS

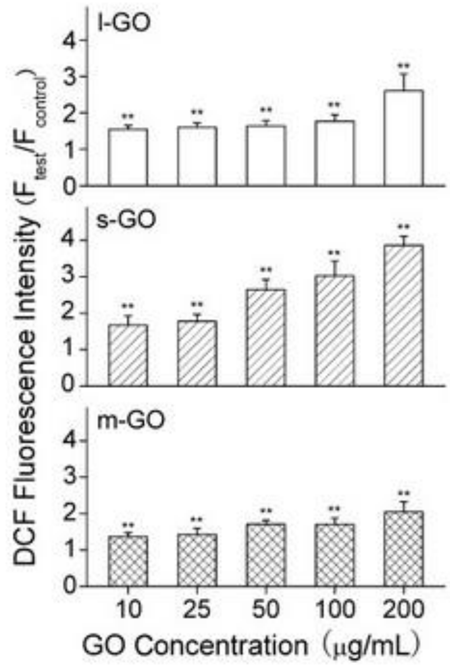

Figure 5: A549 ROS ratio vs. GO concentration.

are DCF fluorescence intensity, which is ultimately the ratio of ROS production by the test group vs. the ROS production by the control. The results show that there is indeed ROS production, most of which can be noticed significantly between 50 and 100 micrograms per milliliter. While ROS ratio follows a roughly linear progression with increasing GO concentration, the actual amounts of ROS production follow a higher order progression.

7.2—Pristine Graphene/ Functionalized Graphene Comparison

The other study that will be considered is comparing pristine graphene (*p*-G)—a sheet of pure carbon atoms—with carboxyl-functionalized graphene (*f*-G)—

graphene with carboxyl functional groups attached—reacted with Vero cells (*17*). The biggest difference between the two types of graphene is that pristine graphene is very hydrophobic and functionalized graphene is hydrophilic, shown below in Figure 6.

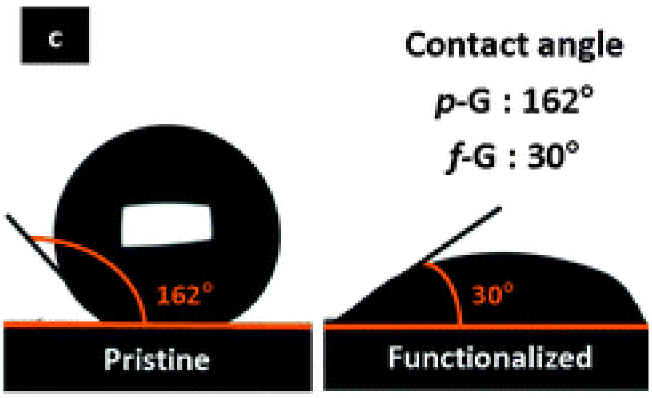

Figure 6: Contact angle comparison of *p*-G and *f*-G.

The hydrophobicity of *p*-G is a key factor in what makes it more toxic than *f*-G because in general, hydrophobicity leads to greater deposition, which causes oxidative stress—essentially the accumulation of ROS—leading to apoptosis. The visualization of the difference between hydrophobic and hydrophilic reactions can be seen below in Figure 7.

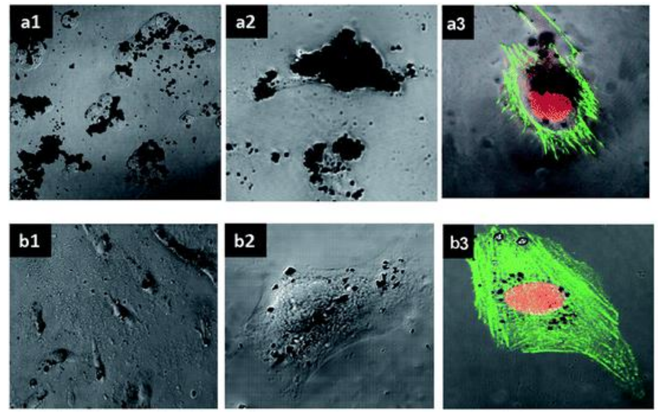

Figure 7: *p*-G and *f*-G differential uptake in Vero cells.

The images on the top are cells treated with pristine graphene, the images on the bottom are cells treated with functionalized graphene. The first two sets of images on the left show how the graphene compounds react with the surface of the cells. One can see the heavy black spots on top show a high amount of deposition caused by the hydrophobicity of pristine graphene, whereas the functionalized graphene does not get deposited, rather it is taken up by the Vero cells, which is much less toxic for them. The third colorful images show the F-actin stabilization for each graphene treatment. High F-actin stabilization indicates that the structure of the cell is maintained, so lower stabilization indicates that the *p*-G is toxic to the Vero cell. Further confirmation of the toxicity of the p-G relative to f-G is shown in Figure 8 below:

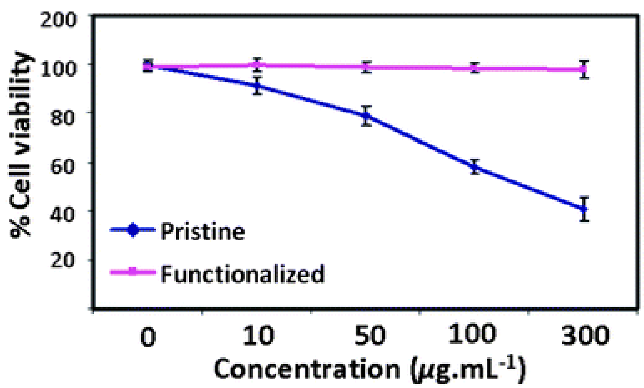

Figure 8: Vero cell viability vs. p-G and f-G concentration. Vero cell viability drops by nearly 50% at high concentrations of *p*-G, while the cell viability remains nearly constant with *f*-G. This is most likely caused by the hydrophobic surface reaction of *p*-G with the Vero cell, causing oxidative stress and therefore apoptosis.

What can be concluded from these studies and from many others is that the most common toxicity mechanism is oxidative stress caused by ROS production, though these effects can be ameliorated by the using functionalized groups on the graphene sheet surface.

8. Conclusion

The modern world depends heavily on computing devices made possible by the transistor. As the world has grown and the demand for computers has grown, the traditional transistor has shrunk to meet this demand. But as we approach the limits of the scaling potential of traditional transistors, and watch energy consumption skyrocket as a result, we must turn to other strategies or give up on increasing computing power. Research has shown that Graphene Nanotransistors could be part of the answer. Using Graphene Nanotransistors can result in faster circuits that also use less power, and are an emerging technology also expected to dramatically improve. Graphene has some concerns as to environmental safety due to exposure at the end-of-life of graphene transistor components, and much still remains to be researched in this realm. However due to the incredible boon Graphene Nanotransistors afford, we believe it well worth the while to continue Graphene research full force simultaneously in their environmental application and their environmental impact.

Works Cited

1. Brown, D. J.; Reams , C. Toward Energy-Efficient Computing. *Queue* **2010,** *8* (2).
2. Cook, G.; Van Horn, J. New Greenpeace report digs up the dirt on Internet data centres, 2011. Greenpeace International. http://www.greenpeace.org/international/en/news/features/New-Greenpeace-report-digs-up-the-dirt-on-Internet-data-centres/ (accessed June 15, 2012).
3. Barroso, L. ; Hölzle, U. The Case for Energy-Proportional Computing. *IEEE Computer* **2007,** *40*, 33-37.
4. Fahey, J. Google guzzles 2.3b kilowatt-hours of electricity a year - enough to power 207,000 homes. http://www.smh.com.au/technology/technology-news/google-guzzles-23b-kilowatthours-of-electricity-a-year--enough-to-power-207000-homes-20110909-1k0l4.html (accessed June 15, 2012).
5. Hawkins, C.; Segura, J. *Introduction to Digital Electronics;* SciTech Publishing: Raleigh, North Carolina, 2010.
6. Markovic, D. Ultralow-Power Design in Near-Threshold Region. **2010,** *98* (2).
7. Schwierz, F. Graphene Transistors. **2010,** *5*.
8. Chen, J.-H.; Jang, C.; Xiao, S.; Ishigami, M.; Fuhrer, M. S. Intrinsic and extrinsic performance limits of graphene devices on SiO2. **2008,** *3* (4).
9. Duan, X. High-speed graphene transistors with a self-aligned nanowire gate. **2010,** *467*.
10. Apple- MacBook Pro Technical Specifications. http://www.apple.com/macbook-pro/specs/.
11. Saxena, V. MOSFET Short Channel Effects. www.lumerink.com/courses/ece5410/Handouts/MOSFET_SCE.pdf.
12. Li, X.; Wang, X.; Zhang, L.; Lee, S.; Dai, H. Chemically Derived, Ultrasmooth Graphene Nanoribbon Semiconductors. **2008,** *319*.
13. Choudhury, M.; Yoon, Y.; Guo, J.; Mohanram, K. Technology exploration for graphene nanoribbon FETs. **2008**.
14. Schwierz, F. Graphene for Electronic Applications -- Transistors and More. **2010**.
15. Aller, B.; Potkonjak, M.; Megeurdichian, S. *Graphene-Silicon Heterogeneity of Transistor Networks;*, Not Yet Published.
16. Ghodrat, M.; Cong, J. Flexible Heterogeneous Simulation Platform. **2011**.
17. Coudert, O. Gate Sizing for Constrained. **1997,** *XX* (Y).
18. Chang, Y. In vitro toxicity evaluation of graphene oxide on A549 cells. **2010**.
19. Sasidharan, A.; Panchakarla, L. S.; Chandran, P.; Menon, D.; Nair, S.; Rao, C. N. R.; Koyakutty, M. Differential nano-bio interactions and toxicity effects of pristine versus functionalized graphene. *Nanoscale* **2011,** *3*, 2461-2464.