Technology Brief 1
Micro- and Nanotechnology

Scale of Things

Our ability as humans to shape and control the environment around us has improved steadily over time, most dramatically in the past 100 years. The degree of control is reflected in the ***scale*** (size) at which objects can be constructed, which is governed by the tools available for constructing them. This refers to the construction of both very large and very small objects. Early tools—such as flint, stone, and metal hunting gear—were on the order of tens of centimeters. Over time, we were able to build ever-smaller and ever-larger tools. The world's largest antenna* is the radio telescope at the Arecibo observatory in Puerto Rico (**Fig. TF1-1**). The dish is 305 m (1000 ft) in diameter and 50 m deep and covers nearly 20 acres. It is built from nearly 40,000 perforated 1 m $\times$ 2 m aluminum plates. On the other end of the size spectrum, some of the smallest antennas today are nanocrescent antennas that are under 100 nm long. These are built by sputtering aluminum against glass beads and then removing the beads to expose crescent-shaped antennas (**Fig. TF1-2**).

Miniaturization continues to move forward: the first hydraulic valves, for example, were a few meters in length (ca. 400 BCE); the first toilet valve was tens of centimeters in size (ca. 1596); and by comparison, the largest dimension in a modern microfluidic valve used in biomedical analysis-chips is less than 100 μm!

The chart in **Fig. TF1-3** displays examples of manmade and natural things whose dimensions fall in the range between 0.1 nm (10^{-10} m) and 1 cm, which encompasses both micrometer (1 μm $= 10^{-6}$ m) and nanometer (1 nm $= 10^{-9}$ m) ranges. ***Microtechnology***, which refers to our ability to manipulate matter at a precision of 1 μm or better, became possible in the 1960s, ushering in an electronics revolution that led to the realization of the laptop computer and the ubiquitous cell phone. It then took another 30 years to improve the manufacturing precision down to the nanometer scale (***nanotechnology***), promising the development of new materials and devices with applications in electronics, medicine, energy, and construction.

*http://www.naic.edu/general/

Figure TF1-1: Arecibo radio telescope.

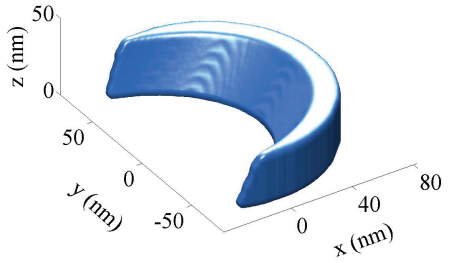
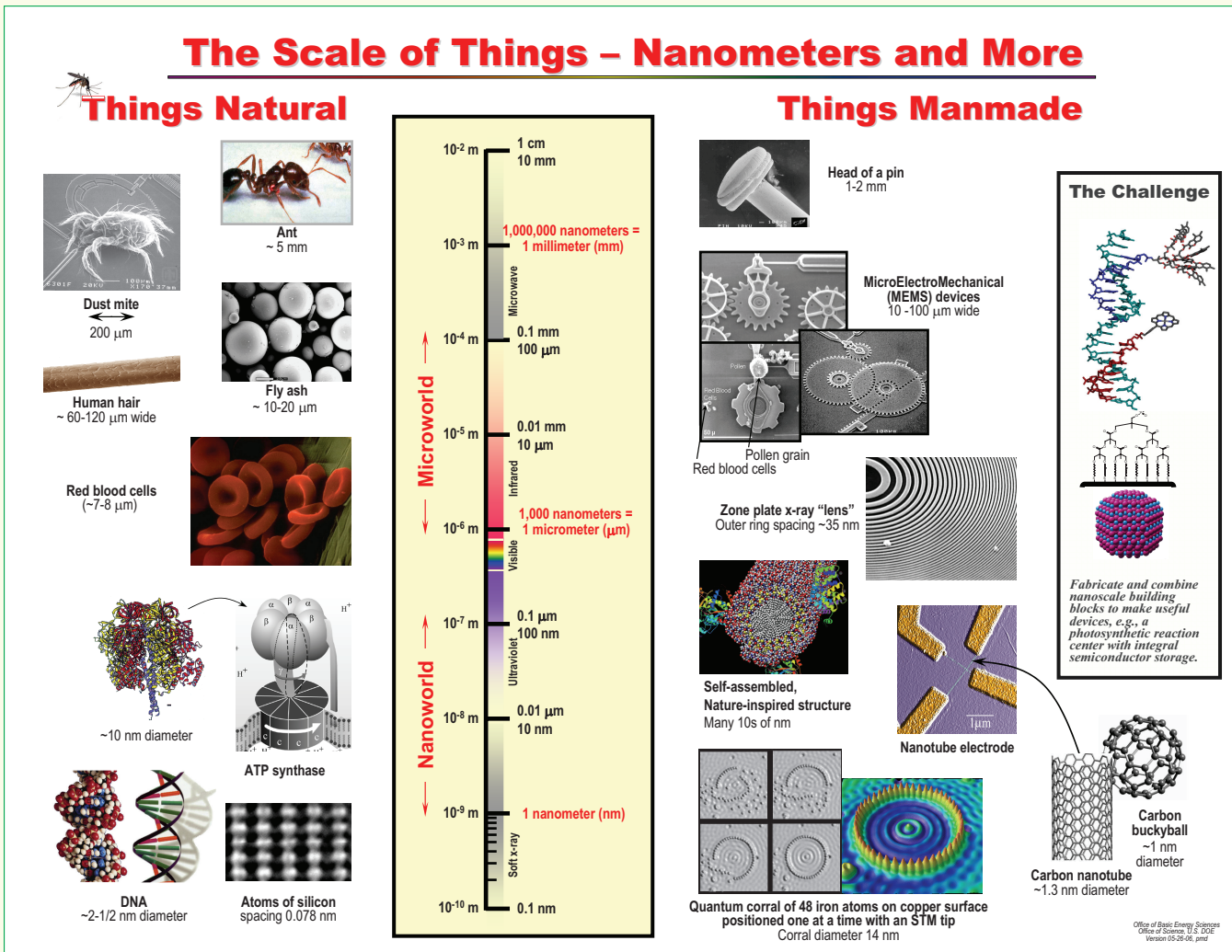


Figure TF1-2: Nano-crescent antenna for use in the ultraviolet range (320 nm to 370 nm wavelength). (Credit: Miguel Rodriguez.)

Moore's Law

With the invention of the semiconductor transistor in 1947 and the subsequent development of the ***integrated circuit*** in 1959, it became possible to build thousands (now trillions) of electronic components onto a single substrate or ***chip***. The 4004 microprocessor chip (ca. 1971) had 2250 transistors and could execute 60,000 instructions per second; each transistor had a "gate" on the order of 10 μm (10^{-5} m). In comparison, the 2006 Intel Core had 151 million transistors with each transistor gate measuring 65 nm (6.5×10^{-8} m); it could



FigureTF1-3: The scale of natural and man-made objects, sized from nanometers to centimeters. (Courtesy of U.S. Department of Energy.)

perform 27 billion instructions per second. The 2011 Intel Core i7 "Gulftown" processors have 1.17 *billion* transistors and can perform ~ 150 billion instructions per second. In recent years, the extreme miniaturization of transistors (the smallest transistor gate in an i7 Core is ~ 32 nanometers wide!) has led to a number of design innovations and trade-offs at the processor level, as devices begin to approach the physical limits of classic semiconductor devices. Among these, the difficulty of dissipating the heat generated by a billion transistors has led to the emergence of ***multicore processors***; these devices distribute the work (and heat) between more than one processor operating simultaneously on the same chip (2 processors on the same chip are called a ***dual core***, 4 processors are called a ***quad core***, etc.). This type of architecture requires additional components to manage computation between processors and has led to the development of new software paradigms to deal with the parallelism inherent in such devices.

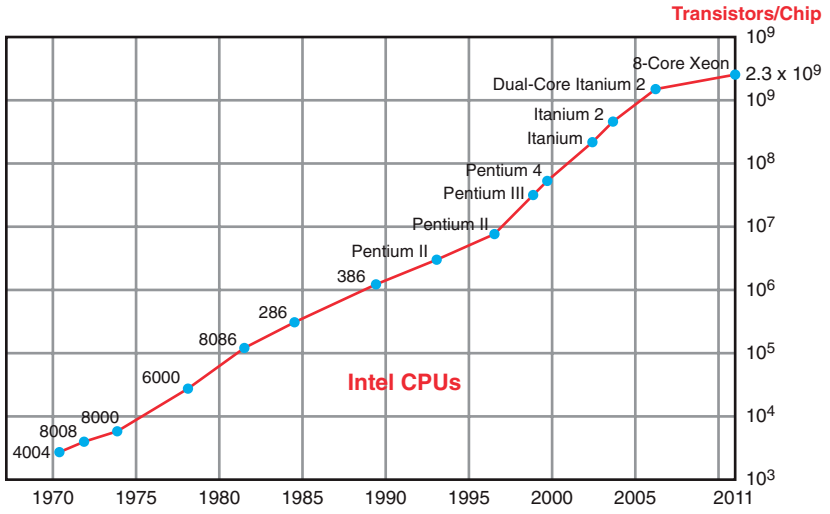


Figure TF1-4: Moore's Law predicts that the number of transistors per processor doubles every two years.

Moore's Law and Scaling

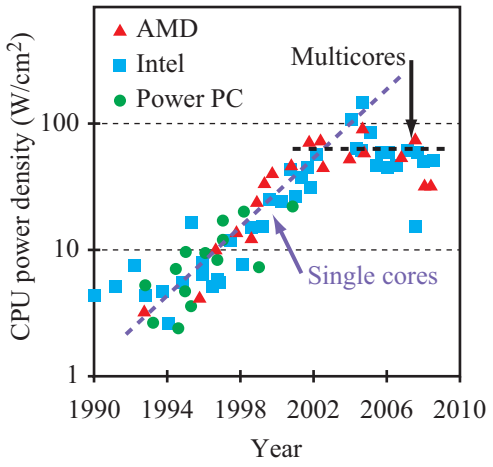
In 1965, Gordon Moore, co-founder of Intel, predicted that the number of transistors in the minimum-cost processor would double every two years (initially, he had guessed they would double every year). Amazingly, this prediction has proven true of semiconductor processors for 40 years, as demonstrated by **Fig. TF1-4**.

In order to understand Moore's Law, we have to understand the basics of how transistors are used in computers. Computers carry all of their information (numbers, letters, sounds, etc.) in coded strings of electrical signals that are either "on" or "off." Each "on" or "off" signal is called a ***bit***, and 8 bits in a row are called a ***byte***. Two bytes are a ***word***, and (when representing numbers) they provide 16-bit precision. Four bytes give 32-bit precision. These bits can be added, subtracted, moved around, etc., by switching each bit individually on or off, so a computer processor can be thought of as a big network of (trillions of) switches. Transistors are the basic switches in computers. We will learn more about them in Chapter 3, but for now, the important thing to know is that they can act as very tiny, very fast, very low power switches. Trillions of transistors are built directly onto a single silicon wafer (read more about how in Technology Brief 7), producing ***very-large-scale integrated*** (VLSI) circuits or ***chips***. Transistors are characterized by their ***feature size***, which is the smallest line width that can be drawn in that VLSI manufacturing process. Larger transistors are used for handling more current (such as in the power distribution system for the chip). Smaller transistors are used where speed and efficiency are critical. The 22 nm processes available today can make lines and features ~22 nm in dimension. They produce transistors that are about 100 nm on a side, switched on and off over 100 billion times a second (it would take you over 2000 years to flip a light switch that many times),[†] and can do about 751 billion operations per watt.[‡] Even smaller, 5 nm transistors are expected to become commercially viable by 2020. The VLSI design engineer uses ***computer-aided design*** (CAD) tools to design chips by combining transistors into larger subsystems (such as logic gates that add/multiply/etc.), choosing the smallest, fastest transistors that can be used for every part of the circuit.

The following questions then arise: How small can we go? What is the fundamental limit to shrinking down the size of a transistor? As we ponder this, we immediately observe that we likely cannot make a transistor smaller than the diameter of one silicon or metal atom (i.e., ~0.2 to 0.8 nm). But is there a limit prior to this? Well, as we shrink transistors down to the point that they are

[†] http://download.intel.com/newsroom/kits/22nm/pdfs/22nm_Fun_Facts.pdf

[‡] https://newsroom.intel.com/servlet/JiveServlet/previewBody/2834-102-1-5130/Intel%20at%20VLSI%20Fact%20Sheet.pdf



	Light Bulb	Integrated Circuit
Power dissipation	100 W	50 W
Surface area	106 cm^2 (bulb surface area)	1.5 cm^2 (die area)
Heat flux	0.9 W/cm^2	33.3 W/cm^2

Figure TF1-5: (a) Plot of heat power density generated by consumer processors over time, (b) comparison of heat power generation of a light bulb with that of a typical processor.

made of just one or a few atomic layers (~1 to 5 nm), we run into issues related to the stochastic nature of quantum physics. At these scales, the random motion of electrons between both physical space and energy levels becomes significant with respect to the size of the transistor, and we start to get spurious or random signals in the circuit. There are even more subtle problems related to the statistics of yield. If a certain piece of a transistor contained only 10 atoms, a deviation of just one atom in the device (to a 9-atom or an 11-atom transistor) represents a huge change in the device properties! This would make it increasingly difficult to economically fabricate chips with hundreds of millions of transistors. Additionally, there is an interesting issue of heat generation: Like any dissipative device, each transistor gives off a small amount of heat. But when you add up the heat produced by more than 1 billion transistors, you get a very large number! **Figure TF1-5** compares the power density (due to heat) produced by different processors over time. The heat generated by single core processors increased exponentially until the mid-2000s when power densities began approaching 100 W/cm^2 (in comparison, a nuclear reactor produces about 200 W/cm^2!). The inability to practically dissipate that much heat led, in part, to the development of multicore processors and a leveling off of heat generation for consumer processors.

None of these issues are insurmountable. Challenges simply spur creative people to come up with innovative solutions. Many of these problems will be solved, and in the process, provide engineers (like you) with jobs and opportunities. But, more importantly, the minimum feature size of a processor is not the end goal of innovation: it is the means to it. Innovation seeks simply to make *increasingly powerful* computation, not smaller feature sizes. Hence, the move towards multicore processors. By sharing the workload among various processors (called ***distributed computing***) we increase processor performance while using less energy, generating less heat, and without needing to run at warp speed. So it seems, as we approach ever-smaller features, we simply will creatively transition into new physical technologies and also new computational techniques. As Gordon Moore himself said, "It will not be like we hit a brick wall and stop."

Scaling Trends and Nanotechnology

It is an observable fact that each generation of tools enables the construction of an even newer, smaller, more powerful generation of tools. This is true not just of mechanical devices, but electronic ones as well. Today's high-power processors could not have been designed, much less tested, without the use of previous processors that were employed to draw and simulate the next generation. Two observations can be made in this regard. First, we now have the technology to build tools

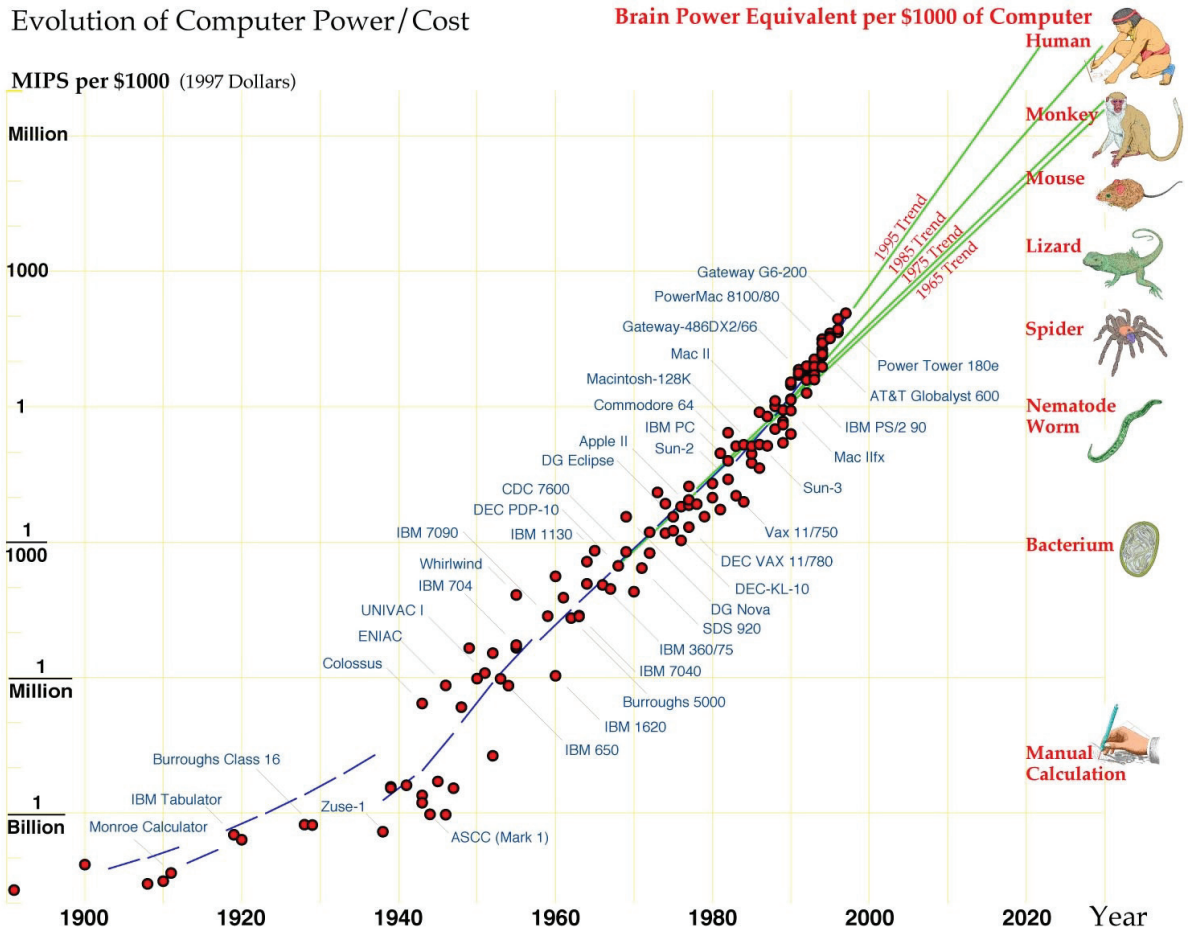


Figure TF1-6: Time plot of computer processing power in MIPS per $1000. (From "When will computer hardware match the human brain?" by Hans Moravec, *Journal of Transhumanism*, Vol. 1, 1998.)

to manipulate the environment at atomic resolution. At least one generation of micro-scale techniques (ranging from ***microelectromechanical systems***—or MEMS—to micro-chemical devices) has been developed that, while useful in themselves, are also enabling the construction of newer, nano-scale devices. These newer devices range from 5 nm transistors to femtoliter (10^{-15}) microfluidic devices that can manipulate single protein molecules. At these scales, the lines between mechanics, electronics, and chemistry begin to blur! It is to these ever-increasing interdisciplinary innovations that the term ***nanotechnology*** rightfully belongs. Second, the rate at which these innovations are occurring seems to be increasing exponentially! (Consider **Fig. TF1-6** and note that the y axis is logarithmic and the plots are very close to straight lines.) Keeping up with rapidly changing technology is one of the exciting and challenging aspects of an engineering career. Electrical engineers use the Institute of Electrical and Electronic Engineers (IEEE) to find professional publications, workshops, and conferences to provide lifelong learning opportunities to stay current and creative (see IEEE.org).