Perspectives

Research directions and challenges in nanoelectronics

R. K. Cavin[1], V. V. Zhirnov[1,3,*], D. J. C. Herr[1], Alba Avila[2] and J. Hutchby[1]
[1]*Semiconductor Research Corporation, Research Triangle Park, Durham, NC, 27709-2053, USA;* [2]*University of the Andes, Carrera 1 Este N. 18 A 70, Bogotá, Colombia;* [3]*Semiconductor Research Corporation, Brighton Hall, Suite 120, 1101 Slater Road, Durham, NC, 27703, USA;* *Author for correspondence (Tel.: +1-919-941-9454; Fax: +1-919-941-9450; E-mail: zhirnov@src.org)*

Abstract

The search for alternate information processing technologies to sustain Moore's Law improvements beyond those attainable by scaling of charge-based devices encompasses several key technologies. Some of these technologies were explored at the Third Workshop on *Silicon Nanoelectronics and Beyond* (SNB III) held at the National Science Foundation in Washington DC in December 2005. They included: (1) non-charge-based devices; (2) devices operated out of thermal equilibrium; (3) alternative interconnect systems; (4) thermal extraction limits and technologies; and (5) fabrication via directed self-assembly. Although this paper was inspired by this highly successful workshop, it is not intended as a summary, but rather an assessment by the authors of some of the fundamental physical considerations evident at the present time.

Introduction

Many research issues pertinent to extending Moore's law scaling for nanoelectronics are being addressed by the portfolio of the U. S. National Nanotechnology Initiative (NNI). A good overview of NNI strategic directions along with emphasis planned for fiscal year 2007 can be found in (Roco, 2006). In January 2005, the National Science Foundation (NSF) and the Semiconductor Research Corporation (SRC) signed a Memorandum of Understanding authorizing the conduct of joint programs in the domain of nanoelectronics. The joint NSF/SRC "Silicon Nanoelectronics and Beyond Program" provides mechanisms for involving industry in NNI research in nanoelectronics. Another example of industry collaboration with NSF is the joint Industry/NSF supplemental funding for six NSF centres focused on discovery of the next switch beyond CMOS.

There remain many formidable challenges to developing technologies for charge-based devices whose performance approaches the best theoretically achievable. For example, the energy consumed per operation by current generation devices is on the order of 10^4 to 10^6 of that which is set by fundamental physics. A broad spectrum of research and development is needed in the underlying disciplines if we are to even approach the physical limits and this will probably include new material systems, new fabrication processes, and new device and interconnect structures. The 2005 edition of the International Technology Roadmap for Semiconductors provides a detailed analysis of technology needs and is a rich resource for research opportunities. It is likely that one-to-two more

decades of intensive research and development will be required to approach the physical performance limits of charge-based electronic systems.

Nevertheless, in view of the long lead times required to transfer research results into useful technologies, there is a need to begin to seek new information processing technologies that will support the continuation of the Moore's Law scaling to extend the exponential benefits in performance-per-unit-cost into the foreseeable future. Several fundamental questions arise when one begin to consider the development of new information processing technologies including

1. Are there physical state variables that could be used for information processing that offer performance advantages relative to charge-based technologies?
2. If we could somehow operate the information processing system when it is out of equilibrium with the thermal bath, could we achieve enhanced performance?
3. How can the performance of interconnect systems for devices be improved?
4. What are the limits of heat removal rates from solids and what technologies can be used to approach these limits?
5. Can directed self-assembly methods abet or supplant conventional assembly technologies for integrated information processing technologies?

In the Section Non-charge-based digital devices, after a brief characterization of the binary switch, we point out that the notion of an energy barrier is useful for the analysis of not only charge-based switches operating in a thermal equilibrium but also can be extended to simple spin and optical binary switches. The energy barrier formalism allows us to offer estimates on the limiting behaviour of these three types of physical realizations for a binary switch. Section Devices operated out of thermal equilibrium is admittedly speculative but it suggests that if switches are allowed to operate out-of equilibrium with the thermal environment, it may be possible to offer improvements in performance relative to their equilibrium-based counterparts. The central idea is that if the switching times of this class of switches can be made substantially shorter than thermal equilibration time, useful and efficient computation may be possible.

Clearly interconnect systems need to be compatible with computational state variables utilizing by the switch to avoid costly conversion. In Section Abstractions for connected binary switches, we focus on the limiting behaviour of electrical interconnects and, using basic arguments on the localization of electrons in a conductor, we find that the energy consumption of limiting charge-based systems is strongly dependent on the number of carriers and on line length. Section Approaching thermal extraction limits presents a basic analysis of heat transfer limits in solids and across interfaces. The developments in this section are intended to give physical insight into those parameters that limit heat extraction and ultimately we show that the 'headroom' for improvement in water cooling greatly exceeds that of air cooling. Section Nanoengineered materials and fabrication via directed self assembly discusses directed self-assembly from an information content perspective and describe possible research opportunities. Section Conclusions provides a brief summary for the paper.

Non-charge-based digital devices

The generic role of energy barriers in binary switches

The basic computational element in known digital information processing systems is binary switch (Figure 1). In its most fundamental form, it consists of:
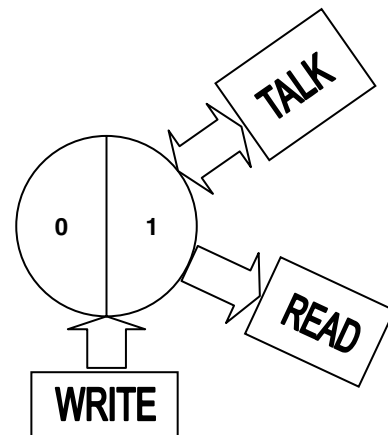

Figure 1. Constituents of an abstract binary switch.

(1) two states 0 and 1 (state variables), which are equally attainable and distinguishable;
(2) a means to control the change of the state (a gate); and
(3) a means to read the state.
(4) a means to communicate with other binary switches.

An important requirement for a binary switch is that it must minimize undesirable spontaneous transitions between binary states (Zhirnov et al., 2003; Cavin et al., 2005a). There are two fundamental sources of errors, which we refer as to Boltzmann errors (thermal noise) and Heisenberg errors (quantum mechanical tunnelling).

The Boltzmann errors occur due to a permanent supply of thermal energy to the system at equilibrium, via mechanical vibrations of atoms (phonons) and via the thermal electromagnetic field of photons (Figure 2).

The existence of random mechanical and electromagnetic stimuli results in a non-zero error probability for the operation of a binary switch. In order to control the error probability, each physical binary switch must contain an energy barrier separating the binary states (Zhirnov et al., 2003; Cavin et al., 2005a). Many critical properties of binary switches, such as size/density, speed, and energy of operation depend on the height and length of the barrier.

The barrier height, E_b, must be large enough to prevent spontaneous transitions (errors). The Bolzmann error occurs when the particle jumps over barrier. This can happen if the thermal energy
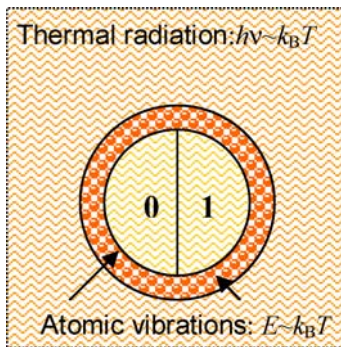
is larger than E_b. The error probability, Π_{err}, is obtained from the Boltzmann distribution as:

$$\Pi_{err1} = \exp\left(-\frac{E_b}{k_B T}\right). \tag{1}$$

The minimum barrier height can be obtained by solving (1) for $\Pi_{err} = 0.5$, the point at which distinguishability of states is completely lost. The result is

$$E_{bmin} = k_B T \ln 2 \approx 0.7 k_B T. \tag{2}$$

The limit (2) was obtained without any assumptions on the physical nature of the system and applies to any physical system in equilibrium with a thermal bath (Figure 2). We will use (2) to estimate below the properties of binary switches operating at lowest possible switching energy.

Another class of errors, manifested in physical binary switches, are the "quantum" or "Heisenberg" errors, occur due to quantum mechanical tunneling through the barrier of finite width. If the barrier becomes too narrow, spontaneous tunneling through the barrier will destroy the binary information. The combined effects of classical and quantum errors have been discussed elsewhere (Zhirnov et al., 2003; Cavin et al., 2005a).

The physical implementation of an energy barrier depends on the choice of the state variable used by the information processing system. The energy barrier creates a local change of the potential energy of a particle from a value U_1 at the generalized coordinate q_1 to a larger value U_2 at the generalized coordinate q_2 as shown in Figure 3a. The difference $\Delta U = U_2 - U_1$ is the barrier height. In a system with an energy barrier, the force exerted on a particle by the barrier is of the form $F = \partial U/\partial q$. A simple illustration of a one dimensional barrier in linear spatial coordinates, x, is shown in Figure 3a. Note, that the spatial energy changes in potential energy require a finite spatial extension ($\Delta x = x_2 - x_1$). This spatial extension defines a minimum dimension of energy barrier, a_{min}: $a_{min} > 2\Delta x$. In this section we consider the physics of barriers for electron charge, electron spin, and optical binary switches.

Energy barrier in charge-based binary switch

For electrons, the basic equation for potential energy is the Poisson equation

Figure 2. A physical system representing binary switch immersed in a thermal bath with two components: radiation and atomic vibrations.
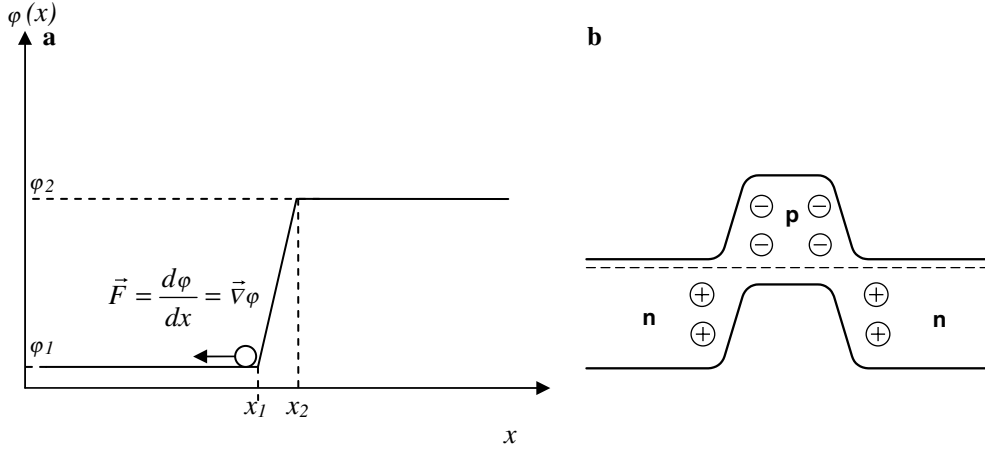
Figure 3. An illustration for the energy barrier in a material system: a – abstraction; b – physical implementation by doping of semiconductor.

$$\nabla^2 \varphi = \frac{\rho}{\varepsilon_0}, \tag{3}$$

where ρ is the charge density, $\varepsilon_0 = 8.85 \times 10^{-12}$ F/m is the permittivity of free space, and ϕ is the potential: $\phi = U/e$. According to (3), the presence of an energy barrier is associated with changes in charge density in the barrier region. The barrier-forming charge is introduced in a material e.g. by doping of semiconductors. This is illustrated in Figure 3b, for a silicon n-p-n structure where the barrier is formed by ionized impurity atoms such as P^+ in the n-region and B^- in the p-region. The barrier height depends on the concentration of the ionized impurity atoms (Sze, 1981)

$$E_b \approx k_B T \ln \frac{N_A^- N_B^+}{n_i}, \tag{4}$$

where N_A^-, N_B^+, and n_i are the concentration of negatively charged impurities (acceptors), positively charged impurities (donors), and the intrinsic carrier concentration in a semiconductor, respectively.

The minimum barrier extension is given by the Debye length (Sze, 1981):

$$L_D \approx \sqrt{\frac{\varepsilon_0 \varepsilon k_B T}{e^2 N}}, \tag{5}$$

where ε is the relative dielectric permittivity of a semiconductor, and $N = N_A^- = N_B^+$ (abrupt p–n junction approximation). The maximum concentration of electrically active dopants N_{max} is close to the density of states in the conduction, N_c, and valence bands, N_v of the semiconductor. For silicon, $N_{max} \sim 10^{19}$ cm^{-3} (Sze, 1981), correspondingly $L_{D\ min} \sim 1.3$ nm. The minimum barrier length therefore is $a_{min} = 2L_{D\ min} \sim 2.6$ nm.

To enable electron movement between these two states, the barrier height must be suppressed from E_b to zero. To do this, the amount of charge in the barrier region needs to be changed. A well-known relation connects the electrical potential difference $\Delta \phi = V$ and charge, Δq, through *capacitance*

$$C = \frac{\Delta q}{\Delta \varphi}. \tag{6}$$

Thus, operation of all charge transport devices involves charging and discharging capacitances to change barrier height thereby controlling charge transport in the device. When a capacitor C is charged from a constant voltage power supply, the energy E_{dis} is dissipated, i.e. converted into heat (Cavin et al., 2005b)

$$E_{dis} = \frac{CV^2}{2}. \tag{7}$$

The minimum energy needed to suppress the barrier (by charging the gate capacitor) is equal to the barrier height E_b. Restoration of the barrier (by discharging gate capacitance) also requires a minimum energy of E_b. Thus the minimum energy required for a full switching cycle is at least $2E_b$.

Next, in order to enable rapid and reliable transition of an electron from state ,0' to state '1', an energy asymmetry between two wells needs to be created. From the distinguishability arguments (Zhirnov et al., 2003; Cavin et al., 2005a), it is easy to show that both the minimum barrier height E_b and the needed energy difference, ΔE_w, between two wells are given by (2). If N is the number of electrons involved in the switching transition between two wells, the total minimum switching energy is

$$E_{SW_{min}} = 2E_b + N\Delta E_w = (N+2)k_B T \ln 2 \quad (8a)$$

If $N = 1$,

$$E_{SW_{min}} = 3k_B T \ln 2 \approx 10^{-20}\,\text{J} \quad (8b)$$

Figure 4 displays ITRS data on the switching energy and gate length of FETs. It also shows the projected scaling trend. Note that the trend line, as an extrapolation of experimental data, ends at a switching energy of $3k_B T$, consistent with (8b). The extrapolated critical dimension of a FET, corresponding to the smallest switching energy, is $\sim$1 nm, in good agreement with theoretical estimates (Zhirnov et al., 2003; Cavin et al., 2005a). Practically, the gate length of FET at the end of scaling is believed to be about 5 nm (ITRS, 2005), not too far from the fundamental limit (Zhirnov et al., 2003; Cavin et al., 2005a). Figure 4 shows that CMOS scaling is close to optimal for electron based devices. Replacement devices must be competitive (size, speed and energy) with charge-based binary switches, which implies operational characteristics better then in charge-based devices.
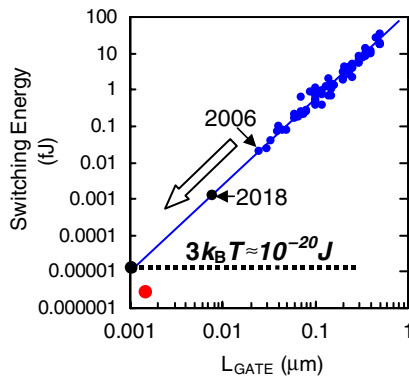

Figure 4. Switching energy trend as a function of MOSFET gate length.

Energy barrier in spin-based binary switch

In addition to charge, e, electrons possess intrinsic angular momentum (spin). As result, they also possess a permanent magnetic moment (Singh, 1997):

$$\mu_s = \pm\frac{1}{2}g \cdot \mu_B, \quad (9)$$

where μ_b is the Bohr magneton, $\mu_B = e\hbar/2m_e$ ($\hbar$ is Planck constant and m_e is electron mass), and g is the coupling constant known as the Landé gyromagnetic factor or g-factor. For free electrons and electrons in isolated atoms $g_0 = 2.00$. In solids, consisting of a large number of atoms the effective g-factor can be different from g_0.

The energy of interaction, $E_{\mu\text{-}B}$, between a magnetic moment $\vec{\mu}$ and a magnetic field $\vec{B}$ is:

$$E_{\mu-B} = -\vec{\mu} \cdot \vec{B}. \quad (10)$$

For the electron spin magnetic moment in a magnetic field applied in the z direction, the energy of interaction takes two values depending of whether the electron spin magnetic moment is aligned or anti-aligned with the magnetic field. From (9) and (10) one can write, assuming $g = 2$

$$\begin{aligned} E_{\uparrow\uparrow} &= -\frac{e\hbar}{2m_e} \cdot B_z, \\ E_{\uparrow\downarrow} &= +\frac{e\hbar}{2m_e} \cdot B_z. \end{aligned} \quad (11)$$

The energy difference between the aligned and anti-aligned states represents the energy barrier in the spin binary switch and is

$$E_b = E_{\uparrow\downarrow} - E_{\uparrow\uparrow} = 2\mu_B B_z. \quad (12)$$

Equations (11) and (12) represent a physical phenomenon known as Zeeman splitting (Singh, 1997). The operation of a single spin binary switch is illustrated in Figure 5. In the absence of an external magnetic field, there is equal probability that the electron has magnetic moment $+\mu_B$ or $-\mu_B$, i. e. the two states are indistinguishable (Figure 5a). When an external magnetic field is applied (Figure 5b, c), the two states are separated. The lower energy state has higher probability of population and it represents the binary state '1' (Figure 5b) or '0' (Figure 5c) in this system. Binary switching occurs when the external magnetic field changes direction as shown in Figure 5b and c.
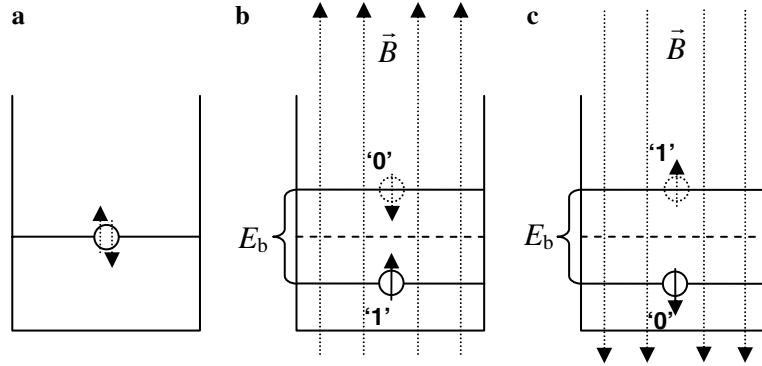
Figure 5. An abstract model of a single spin binary switch: (a) $B = 0$, two states are insdistinguishable; (b, c) $B \neq 0$, two binary states are separated by energy gap E_b.

This abstraction, while very simple, applies to all types of spin devices, at equilibrium with the thermal environment, including, e.g. proposed spin transport devices (Daughton, 1997; Jansen, 2003; Pearton et al., 2005) and coupled spin-polarized quantum dots (Bandyopadhay, 1994).

An energy barrier framework for the spin based switches helps to understand some fundamental constrains of spintronic devices (Zhirnov & Cavin, 2006). The barrier-forming magnetic field B can be either a built-in field formed by a material layer with a permanent magnetization, or created by an external source (e.g. electromagnet). In both cases, the change in the direction of the magnetic field required for binary switching is produced by an electric current pulse in one of two opposite directions. The need for two opposite directions of electrical current requires additional electrical binary switches.

Generic electrical circuits to manipulate the barrier height are a spintronic binary switch is shown in Figure 6. They require four electrical binary switches in the case of single power supply (Figure 6a) or two binary switches in the case of two power supplies (Figure 6b).

Thus each spin-based binary switch needs two or four "servant" charge-based binary switches. This will result in larger area per device and also larger energy consumption per operation, as compared to the charge-based switches. Since the minimum switching energy of one charge-based device is $\sim 3k_B T$ according to (8b), the minimum switching energy of a spin device E_{spin} for single power supply scheme is
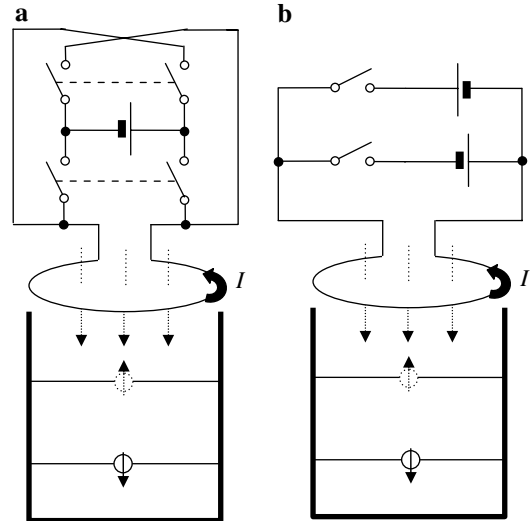

Figure 6. A generic electrical circuit to control spin binary switch: (a) single power supply scheme; (b) double power supply scheme.

$$E_{\text{spin}} = \varepsilon_{\uparrow\downarrow} + 12k_B T, \qquad (13)$$

where $\varepsilon_{\uparrow\downarrow}$ is the "intrinsic" energy to change the spin state. It depends on the specific geometry and the scheme for generating the magnetic field.

One possible direction to address the "electrical" challenge for spin devices is to change the spin control paradigm. The paradigm described above uses a system of binary switches, each of which can be independently controlled by an external stimulus, and each switch can, in principle control any other switch in the system. Can a spin state based device be used to control the state of subsequent

spin devices without going through an electrical switching mechanism as discussed above? Perhaps local interactions can be used to advantage for this purpose (Bandyopadhay, 1994; Cowburn & Welland, 2000). Feasibility assessments of these proposals in general information processing applications are needed.

Let us now consider a hypothetical single spin binary switch that ideally might have atom-scale dimensions. At thermal equilibrium there is a probability of spontaneous transition between spin states '1' and '0' in accordance to (1). Correspondingly, for $\Pi_{err} < 0.5$, according to (2) the energy separation between to state should be larger than $k_B T \ln 2$

$$2\mu_B B_{min} = k_B T \ln 2. \tag{14}$$

From (2) and (12) one can obtain the minimum value of B for a switch operation

$$B_{min} = \frac{k_B T \ln 2}{2\mu_B} = \frac{m_e}{e\hbar} k_B T \ln 2. \tag{15}$$

At $T = 300$ K Eq. (15) results in $B_{min} \approx 155$ T. This is much larger than can be practically achieved (A summary of technologies to generate high magnetic fields is given in Table 1).

One of the most difficult problems of very high magnetic field is the excessive power consumption and Joule heating in electromagnets. The relation between power consumption, P, and magnetic field B of a electromagnet is (Motokawa, 2004):

$$P \sim B^2 \tag{16}$$

and producing magnetic fields e.g. $B > 10$ T requires many mega-watts of power. As result, these magnetic field production systems have large dimensions and a mass of many tons (see Table 1).

Thus, we conclude that *single electron spin devices* operating at room temperature would require local magnetic fields higher than have been achieved to date with large volume apparatus.

In multi-spin systems, it is possible to increase the magnetic moment μ and therefore, to decrease the magnitude of the external magnetic field B required for binary switch operation. The increase of μ can be due to an increase in number of co-aligned spins, which results in collective effects such as paramagnetism and ferromagnetism.

In a system of N spins in an external magnetic field, there are $N_{\uparrow\uparrow}$ spin magnetic moments parallel to the external magnetic field, and the resulting magnetic moment is

$$\mu = \mu_B \cdot N_{\uparrow\uparrow} = \mu_B N(1 - \Pi_{err})$$
$$= \mu_B N\left(1 - \exp\left(-\frac{\mu_B B}{k_B T}\right)\right). \tag{17}$$

As we saw above for all practical cases $\mu_B B \ll k_B T$, and since $(1 - e^x) \approx x$, for $x \to 0$, there results

$$\mu \approx \frac{N\mu_B^2 B}{k_B T}. \tag{18}$$

Equation (18) is known as the Curie law for paramagnetism (Singh, 1997). From (18) and (12) one can calculate the minimum number of electron spins needed for spin binary switch operating at realistic magnitudes of the magnetic field

$$N_{min} \approx \frac{\ln 2}{2}\left(\frac{k_B T}{\mu_B B}\right)^2. \tag{19}$$

For example, for $B = 0.1$ T (small neodymium-iron-boron magnet, see Table 1), $N_{min} \sim 7 \times 10^6$. If the number of electrons with unpaired spins per atom is f (f varies between 1 and 7 for different

Table 1. Examples of practical implementations of the sources of magnetism

Magnet	B	P	Mass	Comments
Small bar magnet	~0.01 T	–	~g	
Small neodymium–iron–boron magnet	~0.2 T	–	~g	
Big magnetic-core electromagnet	~2 T	~100 W	~kg	
Steady-field superconducting electromagnet (Lietzke, 2005)	~16 T	~MW	~tons	Cryogenic temperatures
Current status of Pulse Magnet Program at National High Magnetic Field Laboratory (Marshall et al., 2004)	60–65 T	~MW	~tons	Cryogenic temperatures; a 30 min cooling time between shots; lifetime ~400 cycles

atoms), the number of atoms needed is $N_{\min}/f$. Correspondingly, one can estimate the minimum critical dimension $a_{\min}$ of the binary switch

$$a_{\min} \sim \left(\frac{N_{\min}}{f \cdot n}\right)^{\frac{1}{3}}, \qquad (20)$$

where n is the density of atoms in the material structure. Assuming an atomic density close to the largest known in solids $n \sim 10^{23}$ cm^{-3} and $B \sim 0.1$ T, we obtain $a_{\min} \sim 41$ nm for $f = 1$ and $a_{\min} \sim 22$ nm for $f = 7$. Thus we conclude that for reliable operation at moderate magnetic fields, the physical size of multi-spin based binary switch is larger than the ultimate charge-based devices. The effect of collective spin behavior is currently used in e.g. MRAM and in electron spin resonance (ESR) (a practical sensitivity limit of ESR is $\sim 10^{12}$ spins at room temperature).

Energy barriers in optical binary switch

Optical digital computing was and still is considered by some as a viable option for massive information processing (Wherrett, 1996). Sometimes it is referred as "computing at speed of light" (Higgins, 1995). Indeed, photons move at the speed of light. At the same time, a photon cannot have speed other then the speed of light, c, and therefore it cannot be confined within a binary switch of finite spatial dimensions.

The minimum dimension of optical switch is given by the wave length of light, λ. If $a_{\min} < \lambda$, there is high probability that the light will not 'sense' the state, i.e. the error probability will increase. For visible light, $a_{\min} \sim 400$ nm.

Binary state control in the optical switch is accomplished by local changes in optical properties of the medium, e.g. refraction index, reflectivity, or absorption, while photons are being used to read the state. In many cases, the changes in optical properties are related to a rearrangement of atoms under the influence of electrical, optical or thermal energy. The energy barrier in this case is therefore related to inter-atomic or inter-molecular bonds. Examples of such optical switches are liquid crystal spatial light modulators (Wherrett, 1996) and nonlinear interference filters (Wherrett, 1996). Another approach to energy barriers in optics is to change the refractive index

as result of crystalline-to amorphous phase change, which is used e.g. in the re-writable CD. The minimum switching energy of this class of optical switches is related to the number of atoms, N, and therefore to the size L. In the limiting case, $L \sim a_{\min} \sim \lambda$. In order for atoms of an optical switch to have a distinguishable change of their position, the energy supply to each atom should be larger than $k_B T$. The total switching energy is therefore:

$$E \sim N \cdot k_B T \qquad (21)$$

For minimum energy estimate, we consider the smallest possible N, which corresponds to an single-atom plane of size λ. If optical switch materials have atomic density n, one obtains:

$$E \sim n^{\frac{2}{3}} \cdot \lambda^2 \cdot k_B T \qquad (22)$$

For most solids, $n = 10^{22}$–10^{23} cm^{-3}. Taking $n = 5 \times 10^{22}$ cm^{-3}, $T = 300$ K, and $a_{\min} \sim 400$ nm obtain $E \sim 10^{14}$ J. This is in agreement with estimates of physical limit of switching energy of optical digital switches, given in the literature (Wherrett, 1996).

Another class of optical switches are based on electroabsorption. In these devices, the absorption changes by application of an external electric field that deforms the energy band structure. One example that attracted considerable interest for practical application is the Quantum-Confined Stark Effect (Miller et al., 1984a). If an electrical field is applied to a semiconductor quantum well, the shape of the well is changed (e.g. from rectangular to triangle). As result, the position of energy levels also changes, and this effects optical absorption. Since formation of an electric field requires changes in charge distribution (Eq. 3), the analysis of electroabsorption optical switch is analogous to a charge-based switch, where the energetics is determined by charging and discharging of a capacitor (Eqs. (6) and (7)). Note that the capacitance of optical switch is considerably larger then the capacitance of electron switch because of larger capacitance area of the optical switch ($\sim \lambda^2$). Using the estimated minimum size of an electron switch $a_{e\min} \sim 1$ nm (Zhirnov et al., 2003; Cavin et al., 2005a), and taking into account (9), we obtain an estimate for the switching energy of an electroabsorption device:

$$E \sim 3k_{\mathrm{B}}T\frac{\lambda^2}{a_{e\,\mathrm{min}}^2} \approx 1.2 \cdot 10^{-20}\,\mathrm{J} \cdot \frac{(400\,\mathrm{nm})^2}{(1\,\mathrm{nm})^2}$$
$$\approx 10^{-15}\,\mathrm{J} \tag{23}$$

The result (23) is in an agreement with estimates of physical limit of electro-absorption optical switch, which is given in the literature (Wherrett, 1996; Miller et al., 1984b).

This energy barrier, and therefore the switching energy for optical binary switch is relatively high: the estimate for theoretical limit for optical device switching energy varies between 10^{-14} and 10^{-15} J, for different types of optical switches (Wherrett, 1996). The switching speed is the speed of re-arrangement for atoms or for charge in the material, and it is not related to the speed of light.

Devices operated out of thermal equilibrium

As we have seen from the previous section, an elementary bistable switch can be described at an abstract level in terms of a statistical mechanical system with one degree of freedom and at least two allowed energy states. If the binary switch is in thermal equilibrium, the distribution function characterizing the system will be a Boltzman distribution (1), which results in minimum switching energy $k_{\mathrm{B}}T\ln 2$, (2). At this level of abstraction, all binary switches appear equal and would be governed by the $k_{\mathrm{B}}T\ln 2$ switching limit. However, as we search for alternative ways to physically represent information in future computing systems, it is appropriate to re-examine the critical assumptions in the thermodynamic argument to see if the same level of abstraction is indeed still appropriate. One principle assumptions is that the switch system is at thermal equilibrium with a surrounding thermal reservoir characterized by the temperature T. This implies that there is enough time between switching events to allow the thermal equilibration process to take place.

George Bourianoff (Bourianoff & Nikonov, 2005) pointed out that virtually all the classical literature relating to energy of computation assumed that the logic system had time to thermally equilibrate between switching events. Since the length and time scales for nano electronic devices being considered are the in the same order as thermal scale lengths and times, the fundamental

conclusions relating energy levels to ambient temperature need to be re-examined. In the same context, the presentation pointed out the coupling of ambient thermal energy to magnetic systems (i.e. the effective noise level) was weaker for magnetic systems than charge based systems. The overall conclusion was that non equilibrium systems were very common in nature and consideration of such systems theoretically removed one of the limits associated with classical charge based systems.

It was also pointed out that it is possible to read and write isolated nuclear spin polarized states which maintain themselves in a non-equilibrium energy configuration for time scales on the order of minutes (Datta, 2005). These systems have been experimentally demonstrated in GaAs/AlGaAs systems in the form of the so called "Hyperfine Battery" which reads and writes the nuclear spin stares through the quantum Hall effect (Wurtz et al., 2005).

The concept of devices operating out of equilibrium with the thermal environment has not been yet adequately explored. A detailed analysis of binary switch operation for different state variables in the case where the switch is not at equilibrium with thermal environment is needed to ascertain if there are performance benefits in this case.

Abstractions for connected binary switches

Devices must communicate with one another to support computation, and there is an energy cost associated with communication. In charge based devices this implies that when the electron passes from state '0' to state '1' in one binary device (sending), it needs to control several other devices (receiving). The gates of these receiving devices are electrically coupled to at least one well of the first device (Figure 6). To model this, the length of the sending well is extended to accommodate the gates of receiving devices. In practice, this extension is achieved by interconnect systems. The interconnect length L depends on the number of receiving gates coupled to the line. From simple geometrical considerations, $L > 2Fa$, where a is the barrier width in a binary device and F is number of receiving devices, or fan-out. We now assess how reliably the charge in the extended well of the sending device A in Figure 7 controls the receiving devices B and C. Assume that one electron is

needed to control the barrier of each receiving devices, and one electron passes from one well to another in 0–1 switching ($N = 1$) of the sending device. After device A switched to '1' state, there is one electron on the right-hand well of A.

The electron can freely move along the line of length L and the probability to find it only in the gate B or only in the gate C is given by

$$\Pi_B = \Pi_C = \frac{a}{L} < \frac{a}{4a} = \frac{1}{4} \qquad (24)$$

To increase the probability of successful control, the number of electrons, N, in the interconnect line needs to be increased:

$$\Pi_B = \Pi_C = 1 - \left(1 - \frac{a}{L}\right)^N \qquad (25)$$

For logic operation, a binary switch needs to control at least two other binary switches. The probability that N electrons in the interconnect line of device A will be found in the gates of both B and C is:

$$\Pi_{B \, and C} = \Pi_B \cdot \Pi_C = \Pi_2 = \left(1 - \left(1 - \frac{a}{L}\right)^N\right)^2 \qquad (26)$$

In general,

$$\Pi_n = \left(1 - \left(1 - \frac{a}{L}\right)^N\right)^F \qquad (27)$$

If we require the probability of successful control of all receiving gates to be $> 50\%$, then for $F = 2$ we obtain the minimum number of electrons in the line to be $N_{min} = 5$. For $F = 4$ (typical fan-out in microprocessor ICs), $N_{min} = 14$. Using the obtained minimum number of electrons N_{min}, we can calculate from (8a) the minimum switching energy of device A. The results for minimum ($F = 2$) and typical ($F = 4$) fan-outs are summarized in Table 2 where it is shown that energy dissipation per device in interconnected circuits will considerably exceed isolated device dissipa-
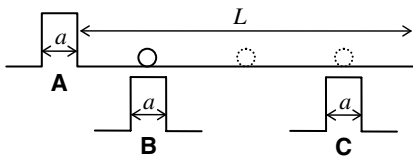
tion. For long distances, the energy trade-offs may favour non-electrical communication technologies (Yablonovich, 2005).

Equations (16)–(19) describe a stationary situation, when the probability Π to find electron in a given point is not function of time. This implies that the charge is already uniformly distributed within the conductor. This model can be extended to include the initial transient processes that determine the velocity of signal propagation.

A schematic energy diagram of a metal interconnect is shown in Figure 8. Before charge q is injected, the conductor is in neutral equilibrium state with minimum electrostatic energy E: $q = 0$ and $E = 0$ (Figure 8a). When charge is injected at one end of the conductor, the electrostatic energy at this end increases and drives the injected electrons to move to the lower-energy part of the conductor with average velocity $\langle v \rangle$ (Figure 8b). The conductor in Figure 8b is in charged non-equilibrium state: $q \neq 0$, $E > 0$, $\langle v \rangle > 0$. After a time interval Δt, the conductor reaches a charged quasi-equilibrium (stationary) state: $q \neq 0$, $E > 0$, $\langle v \rangle = 0$ (Figure 8c). This charged quasi-equilibrium state is the basis for Figure 7 and Eqs. (24)–(27).

The time interval to reach the stationary state can be estimated as:

$$\Delta t \sim \frac{L}{\langle v \rangle} \qquad (28)$$

The average charge velocity can be related to the resistance of the conductor R through the Ohm's Law:

$$I = e \cdot n_e \cdot \langle v \rangle \cdot A = \frac{V}{R} \qquad (29)$$

And therefore,

$$\langle v \rangle \sim \frac{1}{R} \qquad (30)$$

From (28) and (30), we obtain:

$$\Delta t \sim R \cdot C = \alpha \cdot R \cdot L = R \cdot C, \qquad (31)$$

Figure 7. Device abstraction for connected binary switches.

Table 2. Switching energy considering interconnects

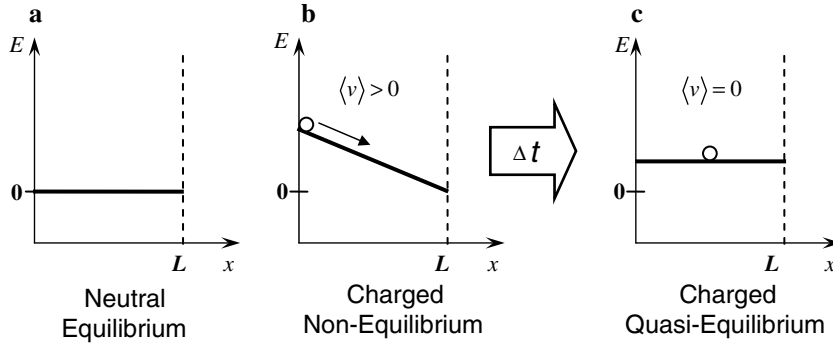| Fan out (n)		2	4
$\Pi_n = 50\%$	N_{min}	5	14
	$E_{SW\ min}$	$7k_B T \ln 2$	$16k_B T \ln 2$
$\Pi_n = 99\%$	N_{min}	17	42
	E_{SWmin}	$19k_B T \ln 2$	$44k_B T \ln 2$

Figure 8. Schematic energy diagram of a interconnect conductor.

where the proportionality constant, α, depends on the geometry of the conductor.

To summarize, increased energy utilization is required to maximize the probability for intended transitions and to minimize the probability of unintended transitions in electronic information processing systems. This is manifested in devices by the need for higher (and wider) energy barriers that must be manipulated to control device operations and by the need to provide a sufficient number of electrons to assure correct operation of downstream devices. Again, by increasing the number of device electrons, there is incurred a corresponding increase in device energy consumption.

Approaching thermal extraction limits

The ability to remove heat has important implications for future computers and, more generally, for nanoelectronic systems. The management of heat is becoming one of the most fundamental problems for the nanoelectonics. In today's computers we deal with heat flows of < 100 W/cm^2, while in a foreseeable future we may need to manage heat flows > 1000 W/cm^2. The purpose of this section is to provide insight into those physical parameters that govern the ultimate limits of heat removal (Zhirnov et al., 2005a, b).

The maximum rate of heat transfer inside an infinite system

The maximum time rate of energy transfer from an atom to a heat transfer agent (e.g. another atom,

electron or photon) is limited by the Heisenberg Principle for time and energy:

$$\Delta E \Delta t \geq \frac{\hbar}{2}, \tag{32}$$

For thermal energy, the maximum $\Delta E \sim k_B T$ and we thus obtain for the maximum rate of energy transfer par channel (*Heisenberg limit*):

$$\dot{Q}_{max} = \frac{\Delta E_{max}}{\Delta t_{min}} = \frac{2(k_B T)^2}{\hbar} \tag{33}$$

If energy is transferred by collisions of masses, the maximum density of the collision channels is the number of atoms per cross-section, say, n_s. Therefore, the rate of heat transfer per unit area is given by:

$$\dot{q}_{max} = n_s \cdot \dot{Q} = \frac{2(k_B T)^2}{\hbar} \cdot n^{\frac{2}{3}} \tag{34}$$

where n is the number of atoms per unit volume.

Equation (34) was obtained under the assumptions that the temperature of the "hot" side of the system $T_1 = 300$ K, while the temperature of the "cold" side $T_2 = 0$ K. If the temperature of the heat receiver $T_2 > 0$, and the heat is transferred through a heterogeneous interface (i.e. atomic/molecular masses of the heat source and heat receiver are not necssarily the same, $m_1 \neq m_2$), the mean energy transfer is given by

$$\langle \Delta E \rangle = 2\mu(E_{1b} - E_{2b})$$
$$= 2\mu(k_B T_1 - k_B T_2) = 2\mu k_B \Delta T \tag{35}$$

where $\mu = \frac{m_1 m_2}{(m_1 + m_2)^2}$

852

For homogeneous medium (equal masses), $\mu = 0.25$, and (33) and (34) will be modified as follows:

$$\dot{Q}_{max} = \frac{(k_B \Delta T)^2}{2\hbar} \tag{36}$$

$$\dot{q}_{max} = \frac{(k_B \Delta T)^2}{2\hbar} \cdot n^{\frac{2}{3}} \tag{37}$$

We now take into account the fact that the interaction time between two neighbouring atoms includes both traverse and contact components. We can obtain a good estimate for the total interaction time from the velocity of sound in the medium, v_s:

$$\Delta t = \frac{l}{v_s} = \frac{n^{-\frac{1}{3}}}{v_s} \tag{38}$$

From (35) and (38) we obtain

$$\dot{Q}_{max} = \frac{\Delta E}{\Delta t} = 2\mu k_B \Delta T v_s n^{\frac{1}{3}}. \tag{39}$$

Assuming $\mu = 0.25$ ($m_1 = m_2$), the heat flow per unit area is

$$\dot{q}_{max} = \dot{Q}_{max} \cdot n^{\frac{2}{3}} = \frac{1}{2} k_B \Delta T v_s n \tag{40}$$

Equations. (39) and (40) represent the maximum rate of heat transfer in the *kinetic limit*, i.e. taking into account the traverse time.

Table 3 compares maximum heat flows calculated in *Heisenberg* and *kinetic* limits for $T_1 = 400$ K and $T_2 = 300$ K. Values for silicon, air and water are presented.

The maximum amount of heat that can be transferred outside a finite system

If we consider heat transfer through the interface between medium 1 with the maximum heat transfer rate $\dot{q}_{1max}$ and medium 2 with the maximum heat transfer rate $\dot{q}_{2max}$, the maximum attainable interface heat transfer rate is

$$\dot{q}_{12max} = \min(\dot{q}_{1max}, \dot{q}_{2max}) \tag{41}$$

For example, in the heat transfer from silicon into air, the maximum heat flux in silicon is 7.6×10^6 W/cm^2, but no more than 650 W/cm^2 can be transferred from silicon into air. Thus, the limiting factor for heat transfer at the solid–gas interface is due to the *discontinuity in density of the conductance channels* because there are several orders of magnitude difference in atomic density. *Thus we conclude that, by transferring heat from a solid to air, the maximum heat transfer rate cannot exceed* ~650 *W/cm^2* (Table 3). Note that this maximum number assumes that in all cases, the hot atoms of the solid collide with "cold" atoms or molecules, which are maintained at $T_2 = 300$ K. In case of solid–water interface, if after collision with a surface atom, the water molecule absorbs additional energy, ΔE, and therefore becomes "warm", with $T_2 > 300$ K. These "warm" water molecules will reduce the energy transfer rate unless removed immediately.

As shown in Table 3, the supply of the heat to the solid–air or solid–water interface is larger than the heat transfer rate in air or in water, because of larger atomic density and larger speed of sound in solids. Thus the temperature of air/water will increase and the cooling efficiency will correspondingly decrease. To increase cooling efficiency, forced liquid cooling could be employed, where the "warm molecule is forcefully replaced by a "fresh" water molecule at $T = 300$ K. In this case, the cooling rate is determined by the supply rate of "cold" water molecules.

For a liquid flow rate (in m/s) u_0 we estimated the maximum heat transfer rate per channel for the forced liquid cooling to be

Table 3. Maximum heat flows inside an infinite system calculated in Heisenberg and kinetic limits for $T_1 = 400$ K and $T_2 = 300$ K

Medium	Speed of Sound, v [cm/s]	Atomic concentration, n [cm^{-3}]	$\dot{q}_{max}$ (Heisenberg limit) [W/cm^2]	$\dot{q}_{max}$ (Kinetic limit) [W/cm^2]
Si	2.2×10^5	5.0×10^{22}	1.2×10^7	7.6×10^6
Water	1.5×10^5	3.4×10^{22}	9.5×10^6	3.5×10^6
Air	3.5×10^4	2.7×10^{19}	8.2×10^4	6.5×10^2

$$\dot{Q} = \frac{\langle \Delta E \rangle}{\Delta t} = \mu \left(2k_{\mathrm{B}}\Delta T - m_2 u_0^2 \right) \cdot u_0 \cdot n_2^{\frac{1}{3}}, \qquad (42)$$

and per unit area

$$\dot{q} = \dot{Q} \cdot n_2^{\frac{2}{3}} = \mu \left(2k_{\mathrm{B}}\Delta T - m_2 u_0^2 \right) \cdot u_0 \cdot n_2. \qquad (43)$$

The maximum of (43) is

$$\dot{q}_{\max} = \frac{4\mu n_2}{3} \sqrt{\frac{2(k_{\mathrm{B}}\Delta T)^3}{3m_2}}, \qquad (44)$$

and is reached at liquid flow rate

$$u_{0\,\max} = \sqrt{\frac{2k_{\mathrm{B}}\Delta T}{3m_2}}. \qquad (45)$$

Figure 9a shows the heat transfer rate $\dot{q}(u_0)$ by forced liquid cooling of silicon by water. The temperature of silicon surface is $T_1 = 400$ K and the temperature of water is $T_2 = 300$ K. For these conditions, $\dot{q}_{\max} = 2.6 \times 10^5$ W/cm^2 and $u_{0\,\max} = 175$ m/s[1]. Equation (44) also applies to forced air cooling. The plot $\dot{q}(u_0)$ for air cooling is shown in Fig. 9b. For this case, $\dot{q}_{\max} = 172$ W/cm^2 and $u_{0\,\max} = 140$ m/s[1].

The permanent presence of cold molecules on the surface can be achieved either by fast replacement or by using molecules with internal degrees of freedom (i.e. higher heat capacity). When thermal energy is transferred to molecules with additional degrees of freedom, the absorbed energy is equally distributed between all degrees of freedom, and only a fraction of total energy goes to translational degrees of freedom and hence increase temperature. Therefore, such molecules can experience several energy transfers before their temperature increase will be remarkable. Water molecules have very high heat capacity and therefore they can longer remain in contact with hot surface before they will need to be replaced due to temperature increase. As result, the delivery velocity u_0 in (43) can be lower. A similar effect occurs during phase change (for example boiling), when breaking cohesive bonds absorb energy without increase in the temperature. Another positive effect of phase change is fast removal of molecules after contact with hot surface. Phase change cooling processes do not impact the maximum heat transfer rate at the interface, but they rather 'improve' the properties of cooling agent and thus help to move closer to the upper bound numbers of Table 3.

To conclude this section, the ability to remove heat from a volume of mass is fundamentally limited by the the maximum amount of heat that can be ultimately transferred to the ambient, which, e. g. to air or water at $T = 300$ K. Figure 10 summarizes the theoretical limits of heat transfer obtained in this section and compares it to the currently experimental results for air (ITRS, 2005) and water (Tuckerman & Pease, 1981) cooling. As follows from this comparison, for the air cooling, the experimentally achieved level is within an order of magnitude of the fundamental limit. However, for the water cooling, a there is significant gap between experimental achievement and theoretical prediction. More research is needed to fully exploit the potential of active water cooling.
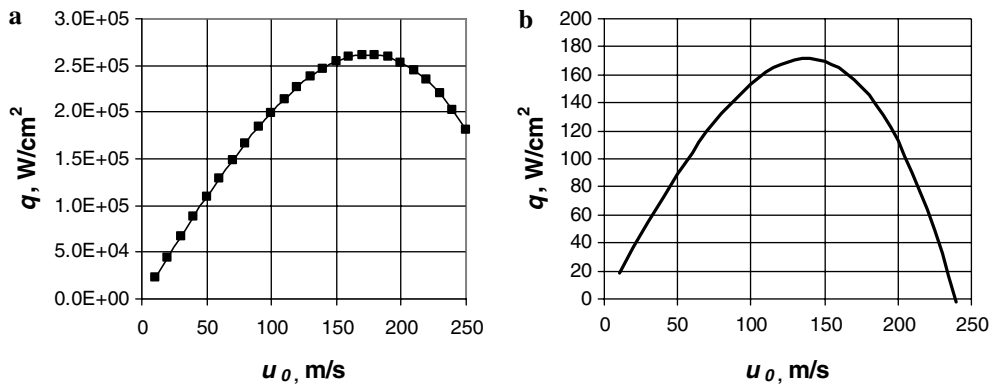

Figure 9. Maximum heat removal rate for forced water (a) and air (b) cooling.
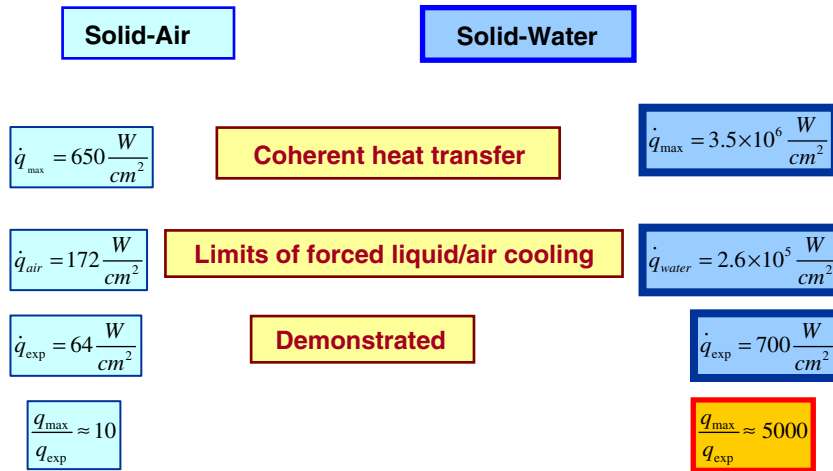
Solid-Air		Solid-Water
$\dot{q}_{max} = 650 \dfrac{W}{cm^2}$	**Coherent heat transfer**	$\dot{q}_{max} = 3.5 \times 10^6 \dfrac{W}{cm^2}$
$\dot{q}_{air} = 172 \dfrac{W}{cm^2}$	**Limits of forced liquid/air cooling**	$\dot{q}_{water} = 2.6 \times 10^5 \dfrac{W}{cm^2}$
$\dot{q}_{exp} = 64 \dfrac{W}{cm^2}$	**Demonstrated**	$\dot{q}_{exp} = 700 \dfrac{W}{cm^2}$
$\dfrac{q_{max}}{q_{exp}} \approx 10$		$\dfrac{q_{max}}{q_{exp}} \approx 5000$

Figure 10. Prediction/Achievement gap for the heat transfer to the ambient.

Nanoengineered materials and fabrication via directed self assembly

Considerable research is needed before new families of nanoengineered materials (NEM) and directed self-assembly (DSA) can be utilized as mainstream integrated circuit manufacturing options. An important research goal is to demonstrate the feasibility of integrating and interconnecting novel engineered nanomaterials and self-assembly methods into IC fabrication processes that enable orders of magnitude increases in functional density. This section focuses on two thematic needs: (1) atomic level design and synthetic control of nanoengineered materials and architectures and (2) the form and content of information designed into self assembling materials.

Atomic level design and synthetic control of nanoengineered materials

By 2020, critical dimensions are projected to be at or less than of 50 ± 5 Å for electronic devices (ITRS, 2005). The ability of modern fabrication tools to offer this degree of dimensional control is limited and auxiliary approaches are needed, including new nanoengineered materials and novel fabrication methods. If achieved, this would enable the realization of a manufacturing capability that exhibits nanometer resolution with atomic level tolerances.

Recent reports (Chen et al., 2005; Herr, 2005) from other fields hint at paths for uncovering potentially breakthrough nanomanufacturing options. For example, a recent study of Pd catalyzed vinyl acetate synthesis reveals that subtle variations in Pd–Pd distances effect dramatic changes in vinyl acetate formation rates (Chen et al., 2005). Figure 11a shows that inter-palladium distances in Au(100), 4.08 Å, more closely resembles that of an optimized catalyst, 3.3 Å, than does Pd in Au(111), at 4.99 Å. Figure 11b indicates that a 0.91 decrease in Pd spacing yields a fourfold increase in vinyl acetate formation rates.

Though the science of designed atomic architectures is in its infancy, this result suggests the impact that atomic level design and synthetic control of nanoengineered materials and architectures would have on the manufacturability of deep nanoscale device components, such as atomically designed catalysts for controlled nanotube growth, deterministically doped channels, etc. Consequently, critical dimensional and tolerance projections at and beyond the end of the Roadmap could become significant if novel materials and manufacturing solutions are discovered that exhibit the required atomic architectural and positional control.

The form and content of information designed into self assembling materials

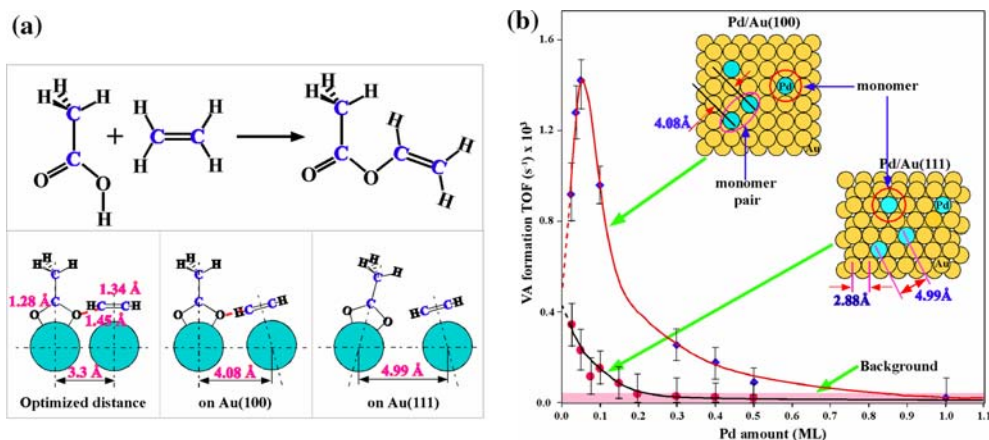Directed self assembling materials, such as block copolymers, exhibit potential as smart resists for

Figure 11. (a) Schematic for vinyl acetate synthesis from acetic acid and ethylene and (b) Vinyl acetate formation rates as a function of Pd coverage on Au(100) and Au(111). From Chen et al. (2005), reprinted with permission from AAAS.

sub-lithographic patterning. The state-of-the-art, with respect to these self assembling systems, can be summarized as follows (Black, 2005a):

- High feature density (small size, tight pitch)
- Periodic structures
- Reasonable size uniformity, little uniformity in ordering
- Some size tenability, but yet to be demonstrated within a single layer
- Robust pattern alignment has yet to be demonstrated

For applications in which many identical elements comprise a single device, there is no need for particularly good order, uniformity, or registration in the element formation. An example of such device is an on-chip decoupling capacitor (Black et al., 2004). In this application, the self-assembled periodic structure can be used to increase surface area. One step up in application complexity is nano-crystal floating gate memory that contains many sublithographic elements (nanocrystals) within gate dielectric (Black, 2005a). In this case, uniformity of size and spacing is important for device reproducibility. Recently, a group at IBM reported fabrication of a Si nanowire FET with several parallel nanowire channels by directed self-assembly of copolymers (Black, 2005b). Critical device dimensions (channel width) were defined by self-assembly within a lithographically defined template.

An important issue in the use of directed SA of block copolymers is the slowness of the SA process, relative to conventional lithographic pro-cesses (Black, 2005a). This question has not been sufficiently discussed by the research community and we believe that the copolymer SA process speed needs to be properly addressed. In fact, quantitative assessments are highly desirable for self-assembly that would characterize (a) informational content delivered to the materials structure; (b) the time required for the SA process, which could be described e.g. as process bandwidth; (c) the energetics of the SA process.

At the simplest level, the information embedded in a materials system is contained in the description of material shapes and their composition (Ayres, 1994). For example, a point in 3D-space is defined by three numbers. A flat surface is completely defined by the four numeric parameters of the form (Ayres, 1994):

$$ax + by + cz + d = 0 \qquad (46)$$

The information required to specify each numeric parameter depends on the precision needed. The information content of a single parametric specification is (Ayres, 1994):

$$I_{par} = \log_2 \frac{m}{\Delta m} \qquad (47)$$

Where m is the dimensional measure and Δm is the required precision. For example, if the specified precision is 1 ppm, $I_{par} = \log_2 10^6 = 20$ bit.

It is also possible to quantify information embedded in materials composition. The number of different atoms available as construction elements of a materials system is given by the number

of elements in the Periodic System, $A \sim 100$. Therefore, the information needed to specify a particular choice of atoms is

$$I_{at} \sim \log_2 A \approx \log_2 100 \approx 7 \text{bit} \qquad (48)$$

Using shape-defining equations of the type (46), along with (47) and (48), it is possible, in principle, to calculate the informational content of an arbitrary materials system. The fabrication process can be regarded as implementing information delivery to a material, e.g. through a series of assembly instructions. As an example, in the case of lithographic patterning, information is delivered by light modulated by mask. In the case of self-assembly methods, more attention needs to be paid to the sources and carriers of information. For example, most current self-assembly methods are based on equilibrium thermodynamics, and as result information is carried only by few process parameters such as temperature, concentration, pressure molecular/atomic composition etc. The resulting materials structure has low informational content, for example a periodic array. The amount of information – the "assembly instructions" – can be considerably increased by using non-equilibrium processes – where some or all of the process parameters (e.g. temperature, concentration etc), are time-dependent, and can be represented as a vector

$$X = \begin{pmatrix} x_1(t) \\ \dots \\ x_n(t) \end{pmatrix} \qquad (49)$$

Careful formulations of information embedded in the material structure as well as the of 'instruction set' used for assembly during fabrication are needed.

Conclusions

The imperative to discover alternative information processing technologies that include binary switches and a means for communication between switches has become an imperative if the exponential improvements in performance per unit cost is to be sustained beyond the limits of charge-based information technology scaling. The urgency associated with this search is caused by the long lead times (~ 12 years) required for technology insertion by the semiconductor industry. In this paper, we have briefly reviewed projected limits for scaling of charge, spin, and optical devices operating in equilibrium with the thermal environment. As a general comment, all of these devices are subject to limits that a multiples of $k_B T$ switching energy dissipation due to the Boltzmann properties of the thermal environment. There is hope that if devices can be operated out-of-equilibrium with the thermal environment, then perhaps computational state variables can be chosen to improve on the $k_B T$ switching energy characteristic of charge-based equilibrium devices. In parallel an aggressive exploration of new architectures is needed to achieve orders of magnitude improvement in effective "function-scalability".

Efficient device communication is required if useful information processing technologies are to be developed. In this paper (and at the workshop), the focus was primarily on the limiting properties of charge-based interconnections. It is argued that in the limit, the interconnect systems will require significantly more energy to operate than the switches themselves. This conclusion does not depend on the metal/dielectric/barrier system used to implement the interconnect system and rests on the localization properties of a few electrons injected into a conductor. If it were possible to eliminate, or reduce the long interconnects, an order of magnitude improvement in energy consumption might be achieved.

Increased component packing densities and higher operating speeds have made higher a real power dissipation inevitable, resulting in the need for enhanced heat management technologies for integrated circuits. This raises the question of physical limits for heat removal; especially for room temperature air and water cooling systems. An intuitively simple model is given that suggests that the relative gains in air cooling capacity that remain to be obtained are quite modest while water cooling continues to offer the capacity for the two to three order of magnitude of improvement.

One of the most costly manufacturing processes is patterning and prospects for cost reductions for top-down methods for nanoscale features are not encouraging. Progress is being made in the use of self-assembly for block co-polymers as resists to achieve increasingly complex features. It appears that line-edge roughness can be improved relative

to top-down patterning by self-assembly methods. IBM has reported experimental use of limited forms of self-assembly in a pre-production facility; however, much research is needed to provide an adequate foundation for the utilization of directed self- assembly in mainstream production of integrated circuits. Perhaps an appropriate goal for the use of self-assembly in nanoelectronic manufacturing is to approach the very high assembly speed, very high complexity and very low energy costs of assembly that is observed in living organisms.

Acknowledgements

References

Ayres R.U., 1994. Information, Entropy, and Progress. AIP Press: New York.

Bandyopadhay S., B. Das & A.E. Miller, 1994. Supercomputing with spin-polarized single electrons in a quantum coupled architecture. Nanotechnology 5, 113–133.

Black C.T., 2005a. Integration of directed self-assembly in semiconductor microelectronics, Third Workshop on Silicon Nanoelectronics and Beyond (SNB), December 8–9, 2005, National Science Foundation, Arlington, VA, USA.

Black C.T., 2005b. Self-aligned self assembly of multi-nanowire silicon field effect transistors. Appl. Phys. Lett. 87: 163116.

Black C.T., K.W. Guarini, Y. Zhang, H.J. Kim, J. Benedict, E. Sikorski, I.V. Babich & K.R. Milkove, 2004. High-capacity, self-assembled metal-oxide-semiconductor decoupling capacitors. IEEE Electron Dev.Lett. 25, 622–624.

Bourianoff G.I. & D.E. Nikonov, 2005. Non-equilibrium systems, alternative state variables and noise considerations. Third Workshop on Silicon Nanoelectronics and Beyond (SNB), December 8–9, 2005, National Science Foundation, Arlington, VA.

Cavin R.K., V.V. Zhirnov, G.I. Bourianoff, J.A. Hutchby, D.J.C. Herr, H.H. Hosack, W.H. Joyner & T.A. Wooldridge, 2005a. A long-term view of research targets in nanoelectronics. J. Nanoparticle Res. 7, 573–586.

Cavin R.K., V.V. Zhirnov, J.A. Hutchby & G.I. Bourianoff, 2005b. Energy barriers, demons, and minimum energy operation of electronic devices. Fluctuation Noise Lett. 5, C29–C38.

Chen M., D. Kumar, C.W. Yi & D.W. Goodman, 2005. The promotional effect of gold in catalysis by palladium-gold. Science 310, 291–293.

Cowburn R.P. & M.E. Welland, 2000. Room temperature magnetic quantum cellular automata. Science 287, 1466–1468.

Datta S., 2005. Non-equilibrium systems: computational state vectors other than charge. Third Workshop on Silicon Nanoelectronics and Beyond (SNB), December 8–9, 2005, National Science Foundation, Arlington, VA, USA.

Daughton J.M., 1997. Magnetic tunneling applied to memory. J. Appl. Phys. 81, 3758.

Herr D., 2005. Novel materials and assembly methods for extending charge-based technology to its ultimate limits, Third Workshop on Silicon Nanoelectronics and Beyond (SNB), December 8–9, 2005, National Science Foundation, Arlington, VA, USA.

Higgins T.V., 1995. Computing with photons at the speed of light. Laser Focus World 31, 72–77.

ITRS, 2005: Semiconductor Industry Association (SIA), International Technology Roadmap for Semiconductors 2005 edition. International SEMATECH, Austin, TX.

Jansen R., 2003. The spin-valve transistor: a review and outlook. J. Phys. D – Appl. Phys. 36, R289–R308.

Lietzke A.F., S.E. Bartlett, P. Bish, S. Caspi, D. Dietrich, P. Ferracin, S.A. Gourlay, A.R. Hafalia, C.R. Hannaford, H. Higley, W. Lau, N. Liggins, S. Mattafirri, M. Nyman, G. Sabbi, R. Scanlan & J. Swanson, 2005. Test results of HD1b, and upgraded 16 Tesla Nb3Sn Dipole Magnet. IEEE Trans. Appl. Supercond. 15, 1123–1127.

Marshall W.S., C.A. Swenson, A. Gavrilin & H.J. Schneider-Muntau, 2004. Development of ''Fast Cool'' pulse magnet coil technology at NHMFL, Physica B 346: 594–598. See also the National High Magnetic Field Laboratory website at: http://www.magnet.fsu.edu/magtech/core.

Miller D.A.B., D.S. Chelma, T.C. Damen, A.C. Gossard, W. Wiegmann, T.H. Wood & C.A. Burrus, 1984a. Band-edge electroabsorption in quantum well structure: the quantum-confined Stark effect. Phys. Rev. Lett. 53, 2173–2176.

Miller D.A.B., D.S. Chelma, T.C. Damen, A.C. Gossard, W. Wiegmann, T.H. Wood & C.A. Burrus, 1984b. Novel hybrid optically bistable switch: the quantum well self-electro-optic effect device. Appl. Phys. Lett. 45, 13–15.

Motokawa M., 2004. Physics in high magnetic fields. Rep. Prog. Phys. 67, 1995–2052.

Pearton S.J., D.P. Norton, R. Frazier, S.Y. Han, C.R. Abernathy & J.M. Zavada, 2005. Spintronics device concepts. IEE Proc.-Circ. Dev. Syst. 152, 312–322.

Roco M.C., 2006. National nanotechnology investment in the FY 2007 budget request. In: AAAS Report on U.S. R&D in FY 2007, Washington, D.C., March 2006.

Singh J., 1997. Quantum Mechanics – Fundamentals and Applications to Technology, John Wiley & Sons.

Sze S.M., 1981. Physics of Semiconductor Devices. John Wiley & Sons.

Tuckerman D.B. & R.F.W. Pease, 1981. Implications of high-performance heat sinking for electron devices. IEEE Trans. Electron Dev. 28, 1230–1231.

Wherrett B.S., 1996. Materials for optical computing. Synthetic Metals 76, 3–9.

Wurtz A., T. Muller, A. Lorke, D. Reuter & A.D. Wieck, 2005. Electrical readout of the local nuclear polarization in the quantum hall effect: A hyperfine battery. Phys. Rev. Lett. 95, 056802.

Yablonovitch E., 2005. The impedance matching predicament in intra–chip communication, Third Workshop on Silicon Nanoelectronics and Beyond (SNB), December 8–9, 2005, National Science Foundation, Arlington, VA, USA.

Zhirnov V.V. & R.K. Cavin, 2006. An energy barrier framework for the limits of charge and spin based switches, J. Nanoelectron. Optoelectron. (to be published).

Zhirnov V.V., R.K. Cavin, J.A. Hutchby & G.I. Bourianoff, 2003. Limits to binary logic switch scaling – a Gedanken model. Proc. IEEE 91, 1934–1939.

Zhirnov V.V., A. Avila & R.K. Cavin, 2005a. Fundamentals of heat removal: limits and opportunities, Third Workshop on Silicon Nanoelectronics and Beyond (SNB), December 8–9, 2005, National Science Foundation, Arlington, VA, USA.

Zhirnov V.V., A. Avila & R.K. Cavin, 2005b. Fundamentals of heat transfer, Internal SRC working paper.