

Electrons that tunnel through insulating barriers and hop on and off minuscule dots are at the heart of future transistor generations

Quantum transistors: toward nanoelectronics

LINDA GEPPERT
Senior Editor

“THE SMALLER WE ARE, the better we perform.” That is the siren song of quantum transistors, in which electrons skip on and off quantum dots or tunnel through barriers thought impenetrable in the world of classical physics.

Nothing stands between these devices and ever-increasing density and performance. Manufacturing processes keep on shrinking their feature sizes, even down to atomic-scale dimensions, and switching frequencies approaching a terahertz are foreseeable because only a handful of electrons is needed to operate the devices.

So great is the allure of quantum devices that several types are under development. One variant is the double-electron-layer tunneling transistor (Deltt) built by researchers at Sandia National Laboratories, in Albuquerque, N.M. Another avenue is to boost the performance of conventional transistors by teaming them with resonant tunnel diodes, quantum devices similar to the Deltts. So far, most resonant tunneling devices (both Deltts and diodes) have utilized indium phosphide or gallium arsenide processes. But engineers are busy building silicon-based devices as well.

Another sort of quantum device shows great promise for nonvolatile memory. Called the single-electron transistor, or sometimes the quantum dot transistor, it is under development by research groups worldwide. A single-electron memory cell—the nanocrystal device pioneered by Sandip Tiwari, now at Cornell University, Ithaca, N.Y.—is of silicon, operates at room temperature, and should prove to have faster read and write times than conventional nonvolatile memories.

Quantum cellular automata are a fourth type of device. Automata are cells that contain four quantum dots arranged in a square. An extra electron resides on each dot of one diagonal or the other, determining if the cell stores a logic 1 or a 0. The cells perform the necessary logic functions by interacting with neighboring cells. The dots can be metal but arrangements of molecules are also possible.

Though quantum transistors are a novelty today, they will be needed once the classical field-effect transistor (FET) can be made no smaller—an event even now on engineers' radar screens. “Already in research labs around the world the last generation of bulk CMOS is being explored,” wrote Hon-Sum Philip Wong in the April 1999 *Proceedings of the IEEE*.

What will eventually stop CMOS technology in its tracks is not the inability to shrink its physical size further, but the dire effect of quantum phenomena on the ever-tinier transistor's operation. In nanoscale FETs, tunneling through ultrathin oxides and extremely narrow channels leaks an unacceptable amount of current. And the minuscule number of dopant atoms in the channel varies enough from one transistor to the next to wreak unavoidable havoc with operating margins [see “The incredible shrinking transistor,” Yuan Taur, *IEEE Spectrum*, July 1999, pp. 25–29].

“Currently we are hoping that we can get down to 20- or 30-nm channel lengths. But below that it looks very difficult to continue with CMOS,” Wong's co-author, David J. Frank, told *Spectrum* recently. Frank is a research staff member at IBM Corp.'s Thomas J. Watson Research Center, in Yorktown Heights, N.Y.

ELECTRONS IN A WELL

Perhaps the most striking quantum effect in transistor-like devices is tunneling. The term alludes to a particle plunging through a barrier that would be impenetrable in the classical world. The mechanism is basic to the Deltt being developed at Sandia.

Though Deltt developers are still in the early stages of exploration, they are pinning their hopes for high speed on performance already obtained with devices that are similar except in having two terminals rather than three. Called resonant tunnel diodes, they have been shown to oscillate at frequencies up to 700 GHz.

Further, Deltts lend themselves to volume manufacturing. They are built with a planar process, using conventional semiconductor deposition. As deposition can be controlled to within a few tenths of a nanometer, it is more than adequate for all structures with dimensions critical to Deltt operation.

Structurally, the Deltt positions an insulating barrier between two two-dimensional wells. In operation, electrons quantum-mechanically tunnel from one well to the other through the barrier [Fig. 1]. Device speed benefits from the tunneling process, which is much faster than the drift of electrons, let alone of slower holes, across a channel. Sandia researchers label the two wells as source and drain, to correspond with conventional transistors.

A well is formed by surrounding one region holding a bevy of free-electron energy states with another region having a dearth of them. An example might be an insulator or a wide-bandgap semiconductor, say, surrounding a metal or narrow-bandgap semiconductor. The well is deemed two-dimensional when it is so narrow in one dimension that electron motion is restricted to the plane of the well.

The dimensions of the well and the height of the barrier (defined as the energy an electron needs to pass over it) determine the discrete energy states allowed to the electrons in the well. An electron can tunnel through the barrier only if its energy and its momentum in the plane of the well are both conserved. In other words, an electron can tunnel only if the energy state of the quantum well

on the other side of the barrier is equal in energy and momentum to the state the electron originally occupied.

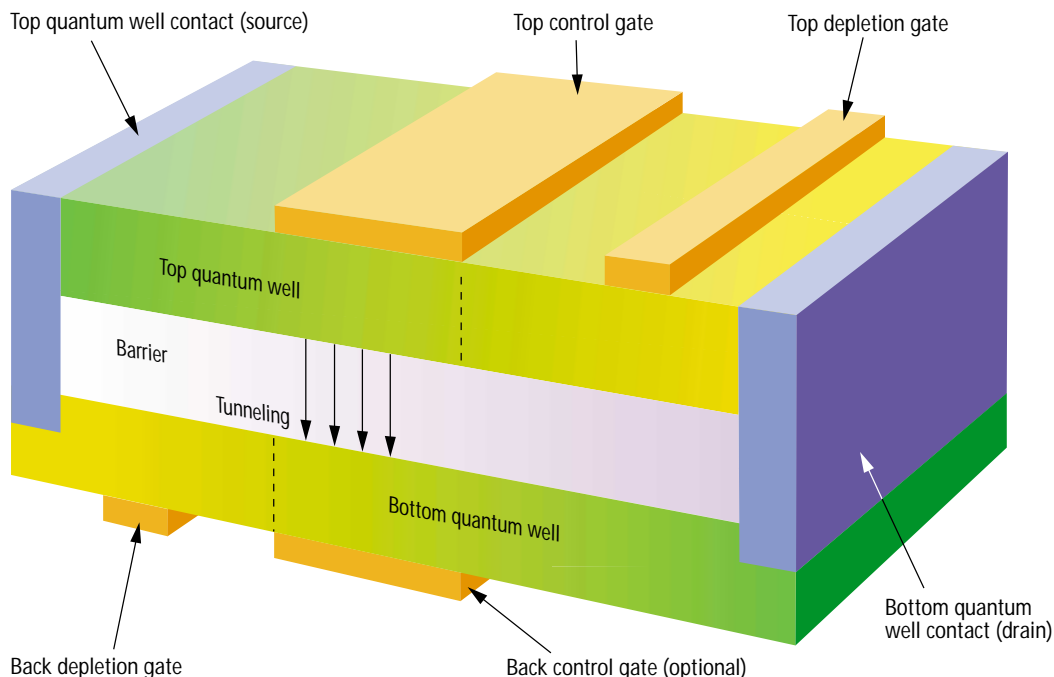
In general, when no voltage is applied to the device, there are no matching states in the two wells, and the device is off. But when the energy level of the electrons in one well is shifted appropriately, the energy states line up opposite each other and tunneling occurs—a condition known as resonance. Applying voltage to a control gate or biasing one of the wells relative to the other will shift the energies. In practice, both voltages are applied in operating the Deltt.

THE BIG DIFFERENCE

In one way, the Deltt's operation is like the switching of an ordinary transistor: at a certain source-drain bias voltage, the device can be switched on or off by applying a voltage to the gate. But there is one important difference. Raise the gate or drain voltage on a conventional transistor, and current increases. But in a Deltt, raising the voltage beyond the point of resonance shuts off the current [Fig. 2]. Put another way, the differential resistance—the change in voltage with respect to a change in current—can become negative. This feature allows complementary circuits to be built with only one type of transistor, rather than the n- and p-type transistors required for CMOS circuits. "It's a unipolar device—the only carriers are electrons. But because the transconductance can take either sign, depending on the control gate voltage, I can make complementary circuits," explained Jerry Simmons, who developed the Deltt. He is Sandia's manager of the semiconductor material and device sciences department.

"That is the exciting thing about these devices," he said, "They are multifunctional, which means that you can perform the same circuit functions with fewer devices." As an example, the Sandia researchers built a static RAM cell using two Deltts in series [Fig. 3]. A conventional CMOS static RAM cell requires n- and p-type transistors, complicating fabrication.

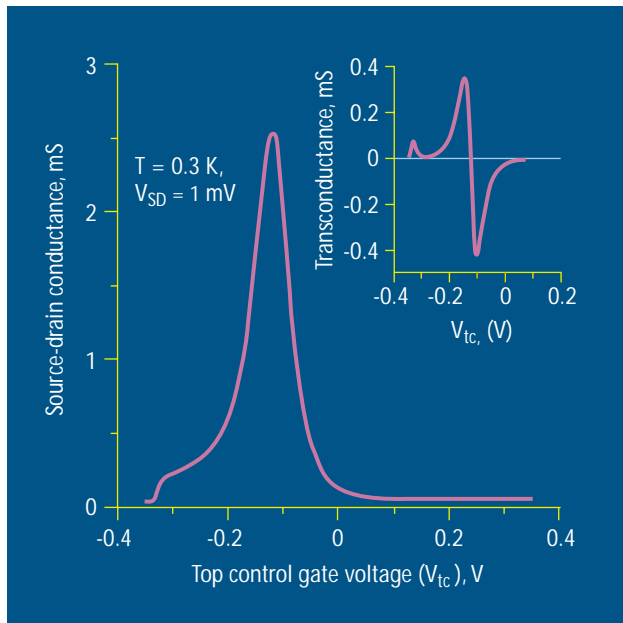
Sandia's Deltt still has a way to go. For one thing, the milli-



[1] The double-layer tunneling transistor, or Deltt, is a planar device consisting of two 15-nm-thick quantum wells separated by a 12.5-nm-thick barrier. Electron tunneling between the wells is controlled by the voltage bias between them and voltage applied to the top control gate. The source and drain ter-

minals contact the top and bottom well, respectively. A top depletion gate rids the top quantum well of electrons in the region next to the drain [yellow shaded area] to ensure that the drain contacts only the bottom quantum well. A back depletion gate does the same for the source terminal.

volt levels at which it operates not only are hard to integrate with the 1-V level of today's electronics but also mean susceptibility to noise. "The final thing," said Simmons, "is that the speed is limited by the Deltt's RC time constant." It is easy to shrink the resistance by using a very thin barrier, but then the two wells are so close that the capacitance is large.



[2] The electrical characteristics of a Deltt at 0.3 K show its resonant behavior. For a constant source-drain voltage, V_{SD} , of 1 mV, the source-drain conductance rises with the voltage on the top control gate. But once past a resonance point (between -0.2 and -0.1 V), conductance drops as the control gate voltage is increased. The transconductance [inset] changes from positive to negative at the resonance voltage.

As a remedy, Simmons and his team have added a third well to the device. In the new configuration, electrons tunnel from the first to the second well, then continue through a second tunnel barrier to the third, very wide well. The large voltage drop across the wide third well allows the electrons to pick up speed.

"We have recently demonstrated such a structure working at voltages on the order of a volt, compatible with existing electronics," said Simmons. "And the source-drain capacitance is much smaller because the starting and ending layers of electrons are very far apart compared to the two wells."

Simmons has also managed to raise the operating temperature. While early devices required temperatures below 4.2 K, the boiling point of liquid helium, more recent Deltts operate at the temperature of ice water—Pepsi temperature, to use his term. The goal, though, is room temperature operation.

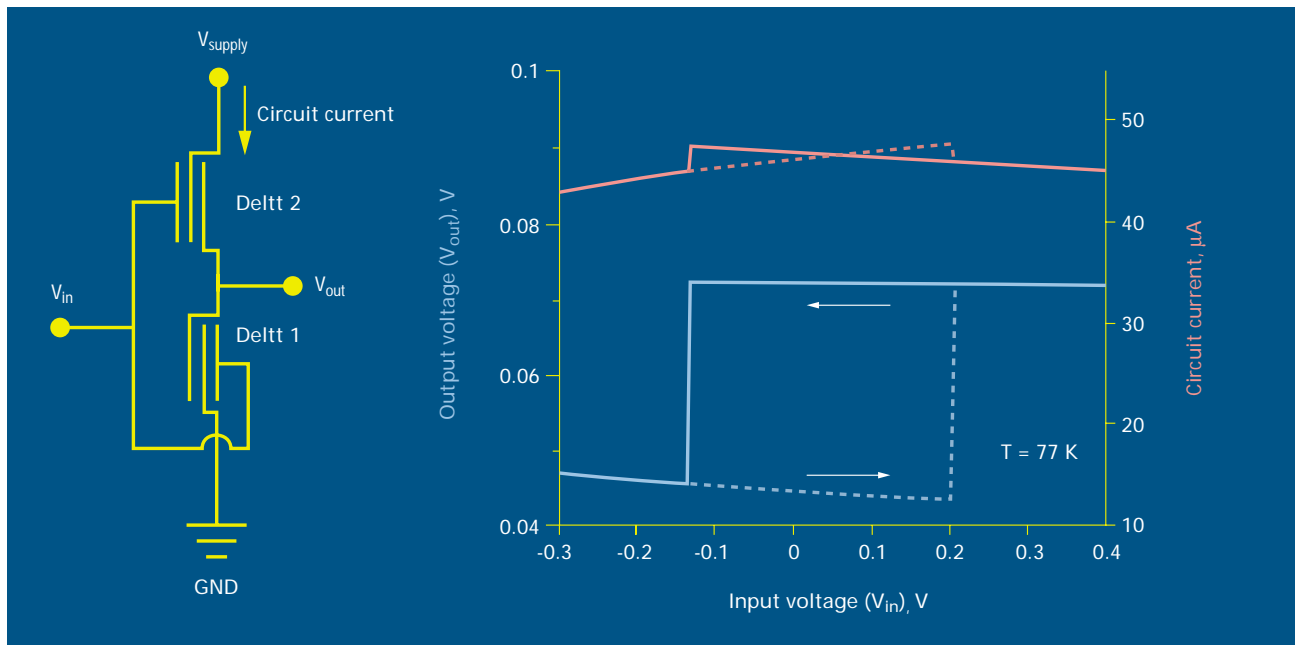
Today the Deltts' relatives, the resonant tunneling diodes (RTDs), routinely operate at room temperature. "But you have to have specialized growth equipment and established infrastructure and expertise to grow systems that will work at room temperature," said Simmons, "and we haven't done it yet." He is working with Jeong-Sun Moon (originally with Sandia's Deltt team) at Hughes Research Laboratories, in Malibu, Calif., to build more advanced structures using indium aluminum arsenide and indium gallium arsenide.

TWO-TERMINAL COUSINS

Although they are not strictly transistors because they lack a third terminal, RTDs are nevertheless finding a role teaming up with ordinary transistors to improve performance of conventional circuits.

In RTDs, electrons tunnel through two barriers separated by a well as they pass from an input source to output drain [Fig. 4]. (These devices differ from Esaki tunnel diodes, which are formed by adjacent regions of heavily doped n- and p-type semiconductor but have no energy well. Instead, electrons tunnel from a state in the conduction band of the n region to an empty state of equal energy in the valence band of the p region.)

Because the RTD well is narrow, the energy levels allowed to electrons are quantized and widely spaced. Typically only one quantized state in the well has any bearing on device operation. When



[3] The Deltt static RAM cell [left] consists of two Deltts connected in series. The same voltage, V_{in} , is applied to the top gate of Deltt 1 and the back control gate of Deltt 2. The supply voltage is 0.12 V. The cur-

rent through the cell, as in a silicon CMOS static RAM cell, is almost constant [curves, top]. A large hysteresis in the output voltage [curves, bottom] demonstrates the bistable properties of the cell.

a voltage is applied between the source and drain, current starts to flow and reaches a maximum at resonance, when the applied voltage raises the energy of the electrons in the source to line up with the well's quantized state. Once that point is passed, current drops. So like the Delts, RTDs display a negative differential resistance.

For this reason “You can put them back to back and they latch,” said Paul Berger, associate professor of electrical engineering at Ohio State University, Columbus. “You can make a static RAM cell with two back-to-back RTDs and one transistor. It is a lot smaller and reduces power consumption.”

In an invited paper at the 1998 International Electron Device Meeting, Alan Seabaugh, professor of electrical engineering at the University of Notre Dame, Ind., and a pioneer of RTD technology, explained the advantages of combining RTDs with conventional transistors to build circuits. They were due, he said, to “the high speed

of the tunnel diode compared to the transistor, and the reduced component count of the tunnel-diode/transistor circuit compared to the transistor circuit.”

For example, if a comparator circuit incorporates RTDs, the number of components is reduced by a factor of six and the area by a factor of four, compared with a circuit built with high-electron-mobility transistors (HEMTs), one of today's fastest devices.

Efforts to develop RTD-based circuits began about 10 years ago at Texas Instruments Inc., Dallas, and moved to Raytheon Systems Co., also in Dallas, when Raytheon purchased Texas Instruments' defense business in 1997. The company is now using RTD-based circuits internally and plans to sample parts to the U.S. Department of Defense and other parties within two years, according to Gary Frazier, manager of the nanoelectronics group at Raytheon.

So far, most work with the RTDs has been done with devices made from GaAs and InP compounds. But a related type of tunnel diode, the silicon resonant interband tunnel diode (RITD), developed by Berger's team, then at the University of Delaware, Newark, promises to increase the speed and shrink the footprints of silicon CMOS circuits also. Preliminary simulations by Seabaugh, then at Raytheon, and his colleagues, show performance improvements for several circuits. For example, a quantizer designed with RITDs and CMOS transistors is less than one-third the size of a circuit designed with CMOS transistors alone. At the same time, it is more than twice as fast, while cutting active power by a factor of 5.8 and static power by 2.3. Seabaugh is presently looking into processes for building RTDs compatible with CMOS manufacture.

“Now is a nice time for this technology to come to fruition,” said Seabaugh, “because with only a single additional mask, CMOS circuit designers get new design flexibility as the end of device scaling is reached.”

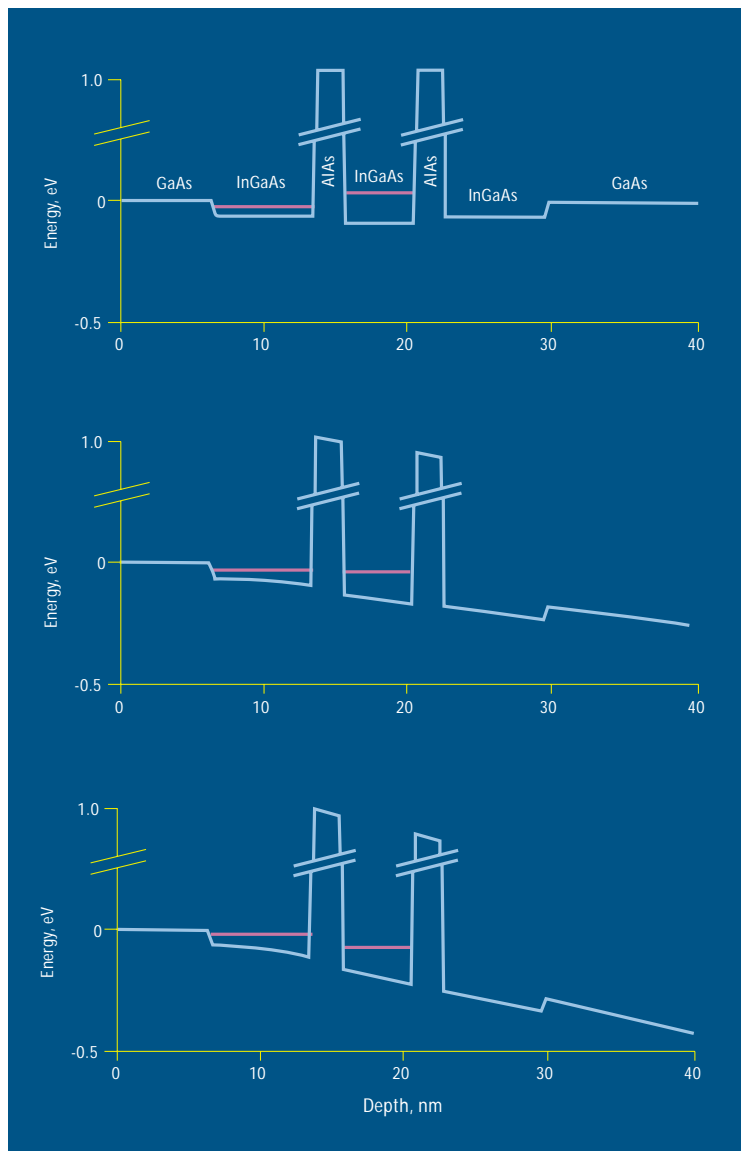
ONE ELECTRON AT A TIME

The basic building block of the single-electron transistor is a small island of conducting material, sometimes called a quantum dot. When the island is small enough, the energy needed to land an electron on it or take one from it depends on how large it is and how many electrons are already on it. For room temperature operation, an island as small as 1–3 nm is needed.

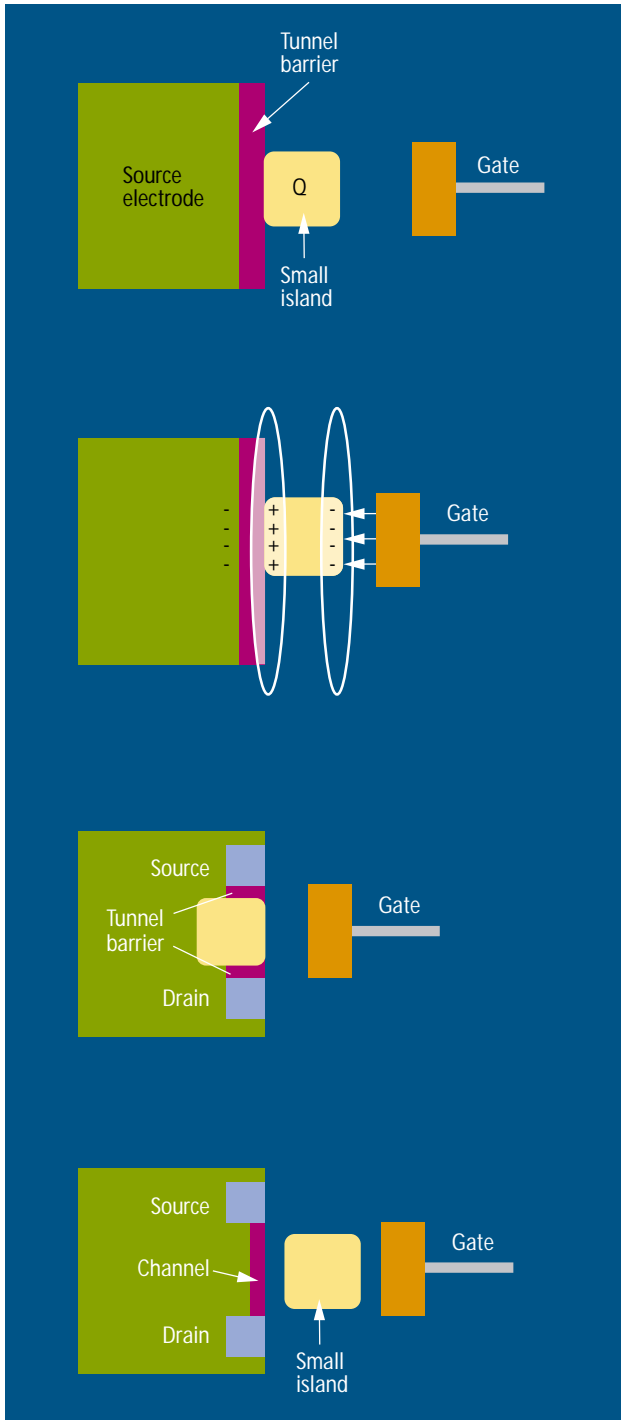
A simple way to get electrons on and off the island is to add an electron source separated from the island by a thin oxide through which electrons can tunnel. A gate over the island changes its energy state, determining the conditions under which electrons tunnel. The result is a structure called a single-electron box [Fig. 5, top].

Applying a gate voltage polarizes the island. At first, as the voltage is increased from zero, an electron in the source lacks enough energy to charge the island. This Coulomb blockade, as it is called, is the basis of all single-electron transistors. But the greater the voltage, the greater the polarization charge becomes, until it equals one electronic charge, whereupon the energy conditions favor the tunneling of one electron to the island from the source electrode [Fig. 5, second from top].

“Such a device can be turned into a transistor,” explained Konstantin Likharev, professor of physics at the University of New York, Stony Brook. One possibility is to replace the channel of an FET by an island and separate it from the source and drain by tunneling barriers [Fig. 5, second from bottom]. As the source-drain voltage is raised, no current flows until a threshold voltage—high enough to overcome the Coulomb blockade—is reached.



[4] In a resonant tunneling diode (RTD), electrons tunnel only when the discrete energy levels in the wells [red lines] are aligned. The left-hand well [top] has too low an energy level for electrons to flow into the central well. A voltage between the device terminals equalizes the levels [center]—the condition for resonant tunneling. Increasing the voltage further raises the left-hand well's energy level above the center well's level, shutting off the current. This RTD was designed by Richard H. Mathews and others at the Lincoln Laboratory of the Massachusetts Institute of Technology.



[5] The basic element in a quantum-dot transistor is called a single-electron box. The dot is a small conducting island separated by a tunnel barrier from a source electrode [top]. A positive voltage on the gate electrode polarizes the charges on the island (though the net charge may be zero), increasing the likelihood that an electron will tunnel between the source and the island [second from top].

One type of quantum-dot transistor replaces the channel of a field-effect transistor (FET) with an island separated by tunnel barriers from source and drain [second from bottom].

A second type of quantum dot transistor, being studied mainly as a nonvolatile memory cell, places the island in the oxide layer between the FET's gate and channel [bottom].

Typically, though, the devices work below the threshold voltage. In this region, the gate voltage controls the source-drain current. "The gate applies a potential to the island and the proper polarity of this potential makes it energetically advantageous for electrons to enter the island," explained Likharev. "So it changes the Coulomb blockade threshold."

As the gate voltage is increased, the blockade threshold voltage drops, and the source-drain current grows until the blockade threshold voltage equals zero. But when the gate voltage is increased beyond this point, the blockade voltage rises again and the current drops.

In other words, the single-electron transistor oscillates between regions of positive and negative transconductance. The advantage is that, like the resonant devices, only one type of device is needed to make a complementary logic gate. "You bias one transistor so that its transconductance is positive and the other so that it is negative, and you get a good analog to CMOS," said Likharev.

Likharev believes that their nanometer dimensions suit single-electron transistors to memory—and possibly some switching applications that he is developing. But they are not suitable for logic. "Because of the small transconductance, it takes a long time to charge up long interconnects," he said. Background charge is also a problem. The sensitive transistor is affected by even a single charged impurity, which occurs often in dielectric materials.

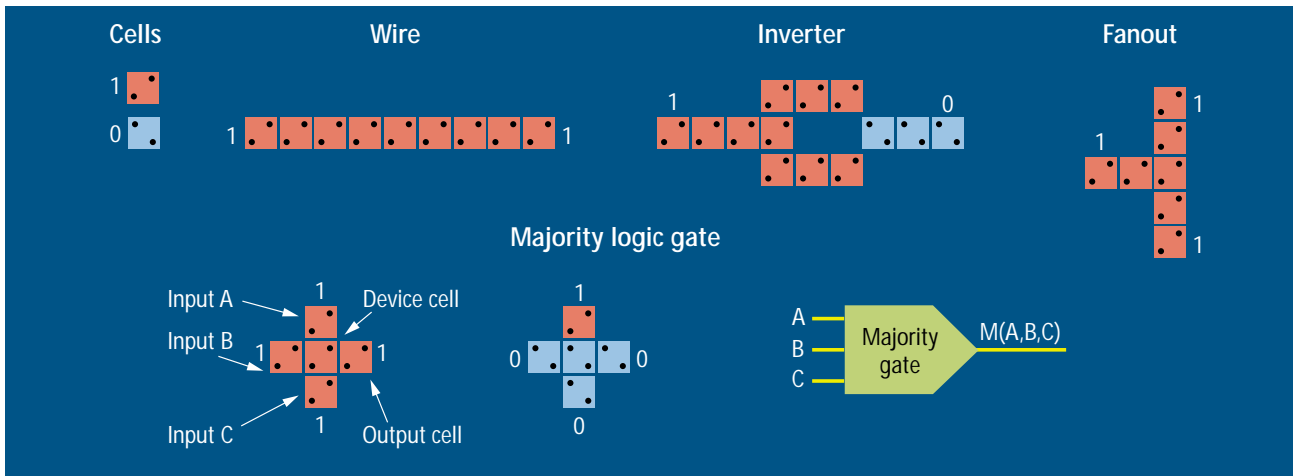
A second important type of single-electron device places the island between the gate and the channel of a field-effect transistor, divided from each by an insulator [Fig. 5, bottom]. The channel becomes the island's electron source, and the transistor gate does double duty as the gate electrode for the island. The combination of the charge stored on the island with the source-drain and gate voltages determines the current through the channel. These devices are so like electrically erasable programmable ROM cells—or flash cells—as to be a natural for non-volatile memory.

In practice, however, fabricating a single nanometer-scale dot and lining it up with a transistor channel is not easy. So some scientists are taking a scattershot approach. They have developed techniques for creating many small dots in the oxide layer between the gate and the channel.

At the Thomas J. Watson Research Center and more recently at Cornell, Sandip Tiwari and his colleagues have formed dots from nanocrystals a few nanometers in diameter, using a chemical vapor deposition process derived from conventional fabrication. The difference, according to Tiwari, a professor of electrical engineering and director of the Cornell Nanofabrication Facility, is that by contrast with the formation of continuous floating gates of conventional flash cells, "you don't allow the grains to coalesce into a continuous film and you allow enough time for the silicon grains to reach a stable state."

The nanocrystal approach has the added benefit of minimizing the effects of so-called interface states—lattice discontinuities that form on the boundaries between the nanocrystals or the gate and the oxide, and that can trap charge. "Even with interface-state densities of mid- 10^{10} or $10^{11}/\text{cm}^2$," said Tiwari, "their effects are overpowered by the fact that there are sufficient nanocrystals around—typically 40 or 50 or more—and with more than one electron stored on them. The effect of those interface states becomes very small."

Another advantage of the nanocrystal structure is that the oxides holding the dots can be much thinner than in conventional flash devices. Just one defect in the continuous floating gate of a conventional flash cell will let electrons escape to the gate or to the overlap regions of the source and drain, Tiwari explained. But a nanocrystalline defect is far from disastrous, for even if the charge on one nanocrystal escapes, charge still exists on all the other nanocrystals. The thinner oxides pay off in write speeds that are faster and voltages that are lower than those of flash cells.



[6] Individual cells in an automata rely on two quantum dots, each with an extra electron [top left]. Repulsion makes the electrons occupy dots on diagonally opposite corners, with logic 1 or 0 assigned according

to the diagonal containing the electrons. Combinations of cells may carry out logic operations. A majority logic AND or OR gate [bottom row] can be built that is a function of inputs A, B, and C [symbol, right].

When the team measured the transistors' current–voltage characteristics, they found that threshold voltages could be increased by fractions of a volt to a few volts if the nanocrystals were charged up. So the presence or absence of charge on the nanocrystals can be used to indicate a logic 1 or 0, discovered by measuring the transistor output at a given gate voltage.

CONNECTING THE DOTS

Rather than assembling a transistor out of quantum dots, Craig Lent and his co-workers at the University of Notre Dame are using quantum dots to perform logic operations. “When you are making things really small, particularly molecular kinds of things, the fundamental idea is not to use the transistor idea at all, but to encode the information in the charge configuration,” Lent told *Spectrum*.

Lent performs computations based not on electron current, but on electron location. The primary building block is a cell of four quantum dots arranged at the corners of a square. Two dots contain one extra electron, and, because of Coulomb repulsion, the two electrons tend to select diagonally opposite dots. A cell with electrons on the upper left and lower right corners could represent a logic 0, and electrons on the opposite diagonal could represent a logic 1 [Fig. 6, top left (only the dots with extra electrons are shown)].

Lent's name for these cells is quantum-dot cellular automata, and for good reason. If two cells are placed close to one another, the electrons in one repel the electrons in the other, both cells automatically assuming the same configuration. So if the first cell is somehow made to be a 1, the one next to it aligns itself to be a 1 also.

Cells can be strung together in a wire to propagate the data in the cell. Better yet, they can be combined to make a family of logic gates consisting of an inverter and a majority gate. Fanout is accomplished with a T-shaped arrangement of cells [Fig. 6, top row].

The majority gate is named for the voting pattern of its three inputs. It is composed of a device cell surrounded by three input cells and an output cell [Fig. 6, bottom row]. It can be viewed as an AND or an OR gate, according to Lent. When the three inputs arrive at the device cell, it has a choice to make. If all three are the same, the cell takes on that value. But if two inputs are a logic 1, it takes on the value of 1, too—the state of the majority. If one of the three is a 1, the device becomes an OR gate of the other two inputs. If one of the three is a 0, the device is an AND of the other two inputs. “So you have everything you need to make logic,” said Lent.

Notre Dame researchers led by Gregory Snider and Gary Bernstein have built majority gates and wires using aluminum dots. Tunnel junctions link each dot to ground and to its two neighbors. An electric field is used to add electrons to a dot or remove them

from it, passing through its tunnel junction with ground. The junctions between dots provide a path for electrons to move among the dots. In more recent devices, the researchers have created electron-hole pairs, with the two electrons on one diagonal and the two holes on the other.

Present devices require extremely cold operating temperatures of around 70 mK. But Lent thinks he can get them to operate at higher temperatures. The operating temperature is keyed to the tunnel junctions. “People have seen them work at room temperature,” said Lent. “We are looking for the best higher-temperature tunnel junction technology we can get.”

But the team's goal is to make the cells out of molecules, enabling them to build circuits with 10^{11} devices per square centimeter. Lent and his team are working with chemists to design and build the molecules. One proposed configuration is a ruthenium-based molecule in which four ruthenium atoms in the four corners of the molecule play the role of the quantum dots. “The d-orbits on the ruthenium atom hold the electron, and it tunnels through a bunch of hydrogen and carbon atoms to get to another ruthenium atom,” explained Lent. So far, the researchers have succeeded in building a double dot—that is, half a cell—and are trying to fabricate the other half.

THE TIME LINE

One-of-a-kind devices of many quantum-transistor varieties have been built and operated. Most need a lot of work before they become practical, particularly for room temperature operation. But even if all the wrinkles are ironed, will the industry take notice?

“CMOS has to reach its limits” in the opinion of IBM's Frank, “before resources and interest will be sufficient to push some other technology far enough to reach some of those application spaces. It's very hard to hit a moving target.” ◆

TO PROBE FURTHER

The April 1999 *Proceedings of the IEEE* has several articles discussing the theory and experimental properties of quantum devices.

The *Digest of the International Electron Devices Meeting*, a compilation of papers presented at the meeting held every December, is a good source for the most recent developments in the technology.

The operation and principles of quantum cellular automata are detailed in two invited papers: “A device architecture for computing with quantum dots,” in the *Proceedings of the IEEE*, April 1997, pp. 541–57, and “Quantum-dot cellular automata: Review and recent experiments,” *Journal of Applied Physics*, Vol. 15, April 1999, pp. 4283–85.