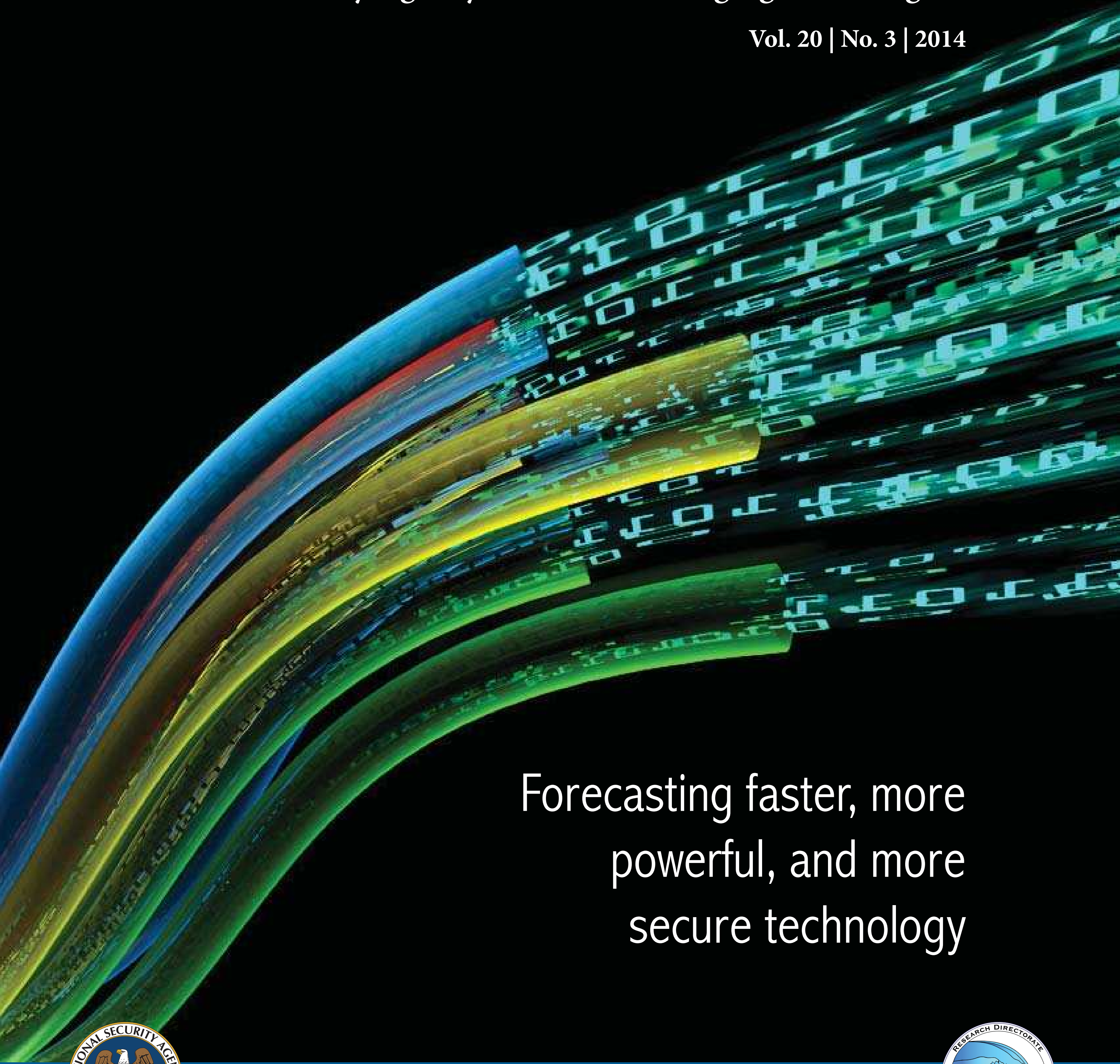
THE Next Wave

The National Security Agency's review of emerging technologies

Vol. 20 | No. 3 | 2014



Forecasting faster, more powerful, and more secure technology

At a Glance | Pointers | Spinouts

THE Next Wave
The National Security Agency's review of emerging technologies

GUEST Editor's column

Communications Technology Forecasting Leader,
Research Directorate, National Security Agency

I would like to thank the staff of *The Next Wave (TNW)* for the opportunity to write this issue's guest editor's column. It is an honor to contribute to *TNW,* especially because this is an issue that looks ahead to a future world in which scientific insights are applied to new or improved technologies that touch our lives. It is in this context that I would like to discuss ***foresight*** and ***the art and science of technology forecasting*** and why these four feature articles are valuable at so many levels.

After deep consideration of a Canadian colleague's clear argument over these past years, I now share his view that foresight is a strategic tool that does use technology forecasting inputs. Furthermore, we agree that foresight is even more than that. Our shared mental model defines foresight as about thinking, debating, and bounding the diverse technology futures that lie ahead. Thus, foresight is the application of critical thinking to long-term developments, trends, and emerging or disruptive technology breakthroughs. Foresight is about anticipating, with adequate lead time, the possibilities. Ultimately, foresight, we believe, informs decisive action.

Foresight activities include

- Examining long-range prospective developments;
- Identifying and understanding key factors and drivers of change;
- Accounting for risk, diversity, and contingencies;
- Anticipating multiple, plausible futures; and
- Highlighting emerging opportunities and threats.

Foresight's contributions to decisive action result in gaming or rehearsal of potential critical challenges and identification of transition strategies that move toward preferred futures.

Drs. Cox and Mosser describe the concept of US Department of Defense (DoD) forecasting which "implies foresight, planning, and careful consideration of how the future operating environment may look" [1]. And they emphasize DoD forecasting implies a "conscious effort to match capabilities to resources" [1]. The authors also note that these activities occur at every level of the defense and security apparatus, and that national policy and strategy are intertwined at the very highest levels.

This approach is reflected in DoD Directive 7024.20 of September 25, 2008, issued by the Deputy Secretary of Defense. Capability portfolio management is described as "optimiz[ing] capability investments across the defense enterprise [so as to] minimize risk in meeting the Department's capability needs in support of strategy" and that this would be done by leveraging the expertise available in various forums and identifying issues, priorities, and capability or resource mismatches for decision makers [2].

The fundamental elements of forecasting and foresight are a) scanning the horizon, b) identifying potentially critical technology, c) predicting the likelihood of emergence, d) anticipating the potentials or effects to business and processes, e) and then optimizing the future capability portfolio in time to remain mission effective. The most difficult problem, of course, is identifying and acting on discontinuous or massively disruptive technologies.

Experiments are under way today that may flatten forecasting and foresight activities in organizations. For example, the Intelligence Advanced Research Projects Activity's Aggregative Contingent Estimation program seeks to "dramatically enhance the accuracy, precision, and timeliness of intelligence forecasts for a broad range of event types" [3]. If successful, the promise seems to be accurate insights and a significant reduction in costs typically associated with full-bore, formal forecasting and foresight activities. One interesting activity within that undertaking is the Good Judgment Project (see http://www.goodjudgmentproject.com).

Similar activities are under way elsewhere. Dreyer and Stang's review of worldwide governmental foresight activities is useful for at least three reasons. First, the reader is presented with a historical review of the foresight movement. Second, key methods are discussed and compared. Third, a number of foresight projects in Australia, New Zealand, the Nordic countries, the European Union, and elsewhere are identified. Implementations in 22 countries are noted [4].

With that said, it is time to turn our attention to the articles and insights of our experts. What are the implications embedded in each of these forecasts? What foresight do we derive from their words?

Communications Technology Forecasting Leader,
Research Directorate, National Security Agency

References

[1] Cox D, Mosser M. "Defense forecasting in theory and in practice: Conceptualizing and teaching the future operating environment." *Small Wars Journal*. 2013;9(1). Available at: http://smallwarsjournal.com/jrnl/art/defense-forecasting-in-theory-and-practice-conceptualizing-and-teaching-the-future-operatin.

[2] England G, US Deputy Secretary of Defense. Department of Defense Directive Number 7045.20: Capability Portfolio Management [accessed 2014 Apr 9]. 2008 Sep 25. Available at: http://www.dtic.mil/whs/directives/corres/pdf/704520p.pdf.

[3] Intelligence Advanced Research Projects Activity. Research Programs: Aggregative Contingent Estimation (ACE) [accessed 2014 Apr 9]. Available at: http://www.iarpa.gov/index.php/research-programs/ace.

[4] Dreyer I, Stang G. "Foresight in governments—practices and trends around the world." In: European Union Institute for Security Studies, *YES 2013: EUISS Yearbook of European Security*. Condé-sur-Noireau (France): Corlet Imprimeur; 2013. p. 7–32.

The Next Wave is published to disseminate technical advancements and research activities in telecommunications and information technologies. Mentions of company names or commercial products do not imply endorsement by the US Government.

This publication is available online at http://www.nsa.gov/research/tnw/index.shtml. For more information, please contact us at TNW@tycho.ncsc.mil.

THE
Next Wave
The National Security Agency's review of emerging technologies

Forecasting superconductive electronics technology

Massachusetts Institute of Technology Lincoln Laboratory

Today's state-of-the-art computer systems are a result of steady, predictable scaling of silicon complementary metal-oxide semiconductor (CMOS) integrated circuit technology. In addition, shrinking transistor dimensions over the past several decades have enabled transistor counts as high as seven billion on commercially available processor chips [1]. However, the energy dissipation of CMOS transistors is reaching physical limits and has become a difficult barrier to building more powerful supercomputers [2]. Advances in "beyond-CMOS" device technologies [3–5] are now seen as a key step towards achieving the next major leap in high-performance computing.

At least an order of magnitude improvement in processor energy efficiency will be necessary before exascale supercomputers are viable. The recently announced Chinese Tianhe-2 machine is reported to operate at a record-breaking 33.9 petaflops [6], where one petaflop is a thousand trillion, or 10^{15}, floating-point operations per second. This performance is achieved by operating close to 80,000 CMOS-based Intel processor chips in parallel. The power consumption of this machine, including the cooling system, is about 24 megawatts. Applying simple scaling, an exascale system providing on the order of 1,000 petaflops would require hundreds of megawatts of power—comparable to a large utility-scale generating station.

One beyond-CMOS technology, digital integrated circuits based on superconductive single-flux-quantum (SFQ) logic, offers a combination of high-speed and ultralow power dissipation unmatched by any other device. First pioneered [7, 8] at Moscow State University in the 1980s by a team led by Professor Konstantin Likharev [9], SFQ technology has seen a resurgence to address the needs of exascale computing.

Operating at cryogenic temperatures, SFQ devices are based on physical phenomena unique to superconductive circuits. Early SFQ research emphasized ultrahigh-speed operation, highlighted by the experimental demonstration reported in 1999 of an SFQ toggle flip-flop operating at an astounding 770 gigahertz (GHz) [10]. Since then, the emphasis for computing applications has shifted towards energy efficiency. Recent advances in SFQ architectures have allowed researchers to develop small-scale, high-speed computational circuits that dissipate more than one thousand times less power than state-of-the-art silicon CMOS circuits—a large energy advantage even after taking into account power for cryogenic cooling. As a result, superconducting SFQ electronics technology may prove advantageous for the future of high-performance computing [11].

Superconductive circuit basics

Superconductivity was first observed by Kamerlingh Onnes in 1911 when he experimentally discovered that the electrical resistance of pure mercury dropped dramatically as the sample temperature dropped below 4.2 kelvins (K) [12]. It was not until the 1950s, following basic advances in quantum mechanics and solid-state physics, that a theoretical basis for superconductivity was developed by physicists Bardeen, Cooper, and Schrieffer [13]. Along with the investigation of the Josephson effect in the early 1960s, these experimental and theoretical breakthroughs set the stage for the technology advances in superconductive SFQ electronics.

At the device level, superconductive SFQ electronics technology provides exceptionally fast, low-energy switching—about 1 picosecond (ps) and 10^{-19} joules (J). In addition, and just as important, SFQ electronics technology offers fast and lossless interconnects between circuit elements. A brief look at some highlights of superconductivity physics provides insight into the technology.

First, superconductors can be described as materials that can carry a direct electrical current (dc) in the absence of an electric field. In other words, the materials have zero resistance at dc. A current can flow between two points in a superconductor without a voltage drop or resistive loss. Niobium, one of the most widely used metals for superconductive electronics, has a critical temperature (T_c) of 9.3 K. Below T_c, niobium is superconductive; above T_c, niobium behaves as a normal metal with electrical resistance.

Superconductors also have very low electrical loss at microwave frequencies. This property enables compact transmission lines for transporting short microwave pulses with minimal energy dissipation at close to the speed of light, in many ways similar to how an optical pulse can be transmitted on optical fiber. This eliminates the capacitive charging energy of interconnects that can dominate the total amount of power dissipation in the most advanced high-speed CMOS circuits.

A second fundamental property of superconductors is flux quantization. The magnetic flux passing through a closed superconducting ring carrying a current is quantized in multiples of the flux quantum expressed simply in terms of fundamental physical constants as $\Phi_0 = h/2e$, about 2 millivolt picoseconds (mV-ps), where h is Planck's constant and e is the electron charge. Superconductive SFQ electronics technology is based on the manipulation and transport of these magnetic flux quanta.

Flux quantization results from the quantum mechanical behavior of metallic superconductors. The theory developed by Bardeen, Cooper, and Schrieffer [13] shows that superconducting electrical current results from electron pairing. These electron pairs, now known as Cooper pairs, combine two electrons, one with spin up and one with spin down. With a net spin of zero, Cooper pairs behave as bosons—this means that at very low temperatures they can all fall, or condense, into the ground state, the lowest energy state of the system. The collection of Cooper pairs can thus be described by a single macroscopic quantum mechanical wave function.

The phase variation of the wave function around a closed superconducting ring must be an integer multiple of 2π, and this leads directly to quantization of magnetic flux. This behavior starkly contrasts that of the single unpaired electrons in conventional conductors, which behave as fermions, a type of particle governed by the Pauli exclusion principle. Since no two fermions can occupy the same quantum state simultaneously, condensation of all conduction electrons into a single ground state is not possible.

Realizing active circuits based on manipulating flux quanta requires a means to switch these quanta in and out of superconducting loops. This is accomplished by interrupting the ring with a Josephson junction (JJ), consisting of two superconductors separated by a thin insulating, or barrier, layer. Quantum mechanics predicts that Cooper pairs can tunnel across the barrier layer. As described by the Josephson effect, the junction superconducting current varies periodically with the phase difference ϕ between the wave function on either side of the barrier as $I = I_c \sin\phi$, where I_c is the critical current.

In addition, a time-varying change in ϕ results in a voltage drop across the junction.

In the steady state, a JJ can support a constant (dc) superconducting current with zero voltage drop and a phase difference ϕ that remains constant over time as long as the current level does not exceed the junction critical current I_c. If the junction current is forced to exceed I_c, then a voltage will develop across the junction along with a time-varying phase ϕ, as indicated in figure 1. Each 2π rotation in ϕ results in the generation of an SFQ voltage pulse of area Φ_0, approximately 2 mV-ps, across the junction. A higher junction voltage corresponds to a faster rate of SFQ pulse generation. The average voltage across a JJ in an SFQ circuit is proportional to the average SFQ switching frequency f by the simple relationship $V = \Phi_0 f$. In actual circuits, a shunt resistor is placed in parallel with the junction to provide damping and well-behaved SFQ pulse generation.

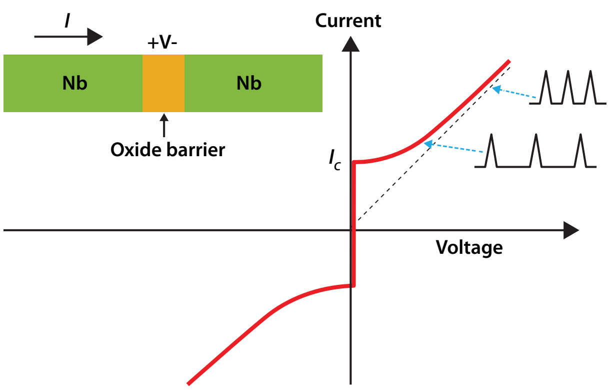


FIGURE 1. This diagram illustrates the current-voltage (*I-V*) relationship for a niobium (Nb)-based Josephson junction (JJ). A drive current in excess of the critical current I_c results in SFQ pulse generation.

Building circuits and logic gates exploiting SFQ operation involves combining loops and inductors for storing flux along with transformers and JJs for control and switching. A very simple SFQ circuit, shown in figure 2, illustrates the basic mechanism. A superconducting ring is interrupted by a single JJ, and a transformer couples an amount of magnetic flux into the ring proportional to an externally applied control current. If the control current results in the loop current I_L exceeding I_c, then a short voltage pulse will result across the junction along with a 2π phase shift. This corresponds to a single quantum of flux passing through the junction.

The basic SFQ switching operation can be extended to form a complete set of logic functions. A simple example of an SFQ gate is the D-type flip-flop, a key building block of SFQ shift registers and shown as a circuit schematic in figure 3. The D flip-flop has a storage loop formed by the Josephson junctions J_1 and J_2 and the inductor L_2. With a bias current applied to keep J_1 close to its critical current, an input 'D' (data) pulse entering through J_0 will switch J_1 and inject an SFQ pulse into the storage loop, resulting in an increase in the circulating current I_s passing through J_2. Readout is performed with an incoming clock pulse. In the presence of a stored pulse I_s, an incoming clock pulse will cause J_2 to switch, resulting in an output pulse at 'Q'. With no stored pulse, the clock pulse is insufficient to switch J_2 and there will be no output pulse at 'Q'.

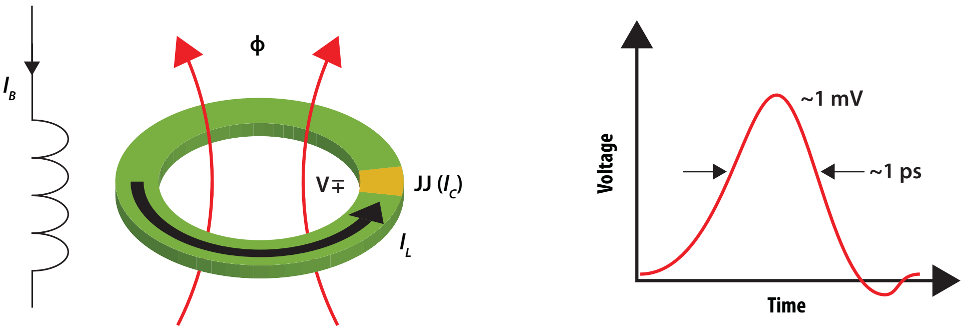


FIGURE 2. This diagram illustrates the generation of an SFQ pulse in a superconducting ring with a Josephson junction (JJ). When the applied current I_B to the transformer results in a circulating loop current I_L in excess of the JJ critical current I_c, an SFQ pulse is generated.

As stated earlier, ultralow energy dissipation is a key attraction of SFQ electronics. The energy dissipated for each basic SFQ switching event is given by the simple expression $\Phi_0 I_c$. For a typical critical current of 50 microamperes (µA), the switching energy is an exceptionally small 1 x 10^{-19} J.

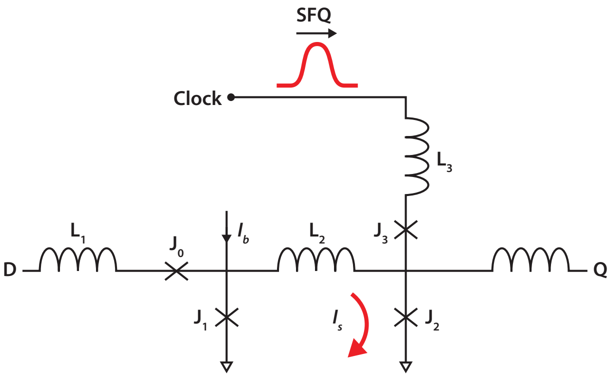


FIGURE 3. This schematic circuit diagram shows an SFQ D-type flip-flop. The circuit is comprised of Josephson junctions (J_n) forming circuit loops with inductors (L_n). A bias current I_b is applied to the storage loop. A 'data' input pulse at D results in a stored SFQ pulse which, in the presence of a clock pulse, is transferred to the output Q.

The switching speed varies in proportion to the square root of the JJ critical current density J_c. For niobium-based technology, a typical J_c value of 10 kiloamperes per square centimeter (kA/cm^2), which provides an I_c of 50 μA for a junction of area 0.5 μm^2, results in an SFQ pulse width of only about 1 ps and a maximum gate clock rate of about 350 GHz for small-scale circuits.

Lowering I_c would be a straightforward way to achieve even lower switching energy, but for practical conventional logic circuits where low error rates are required, it is important to operate with switching energies several orders of magnitude higher than the thermal energy (k_BT), which at 4 K is about 6 x 10^{-23} J. Just as for CMOS logic circuits which contain multiple switching transistors per logic gate, an SFQ logic operation will require multiple SFQ switching events. With a typical SFQ logic gate configured with roughly five JJs, gate switching is still exceptionally low at about 5 x 10^{-19} J.

To put SFQ switching energy and speed in perspective, comparisons with other advanced high-speed logic technologies are shown in figure 4. The minimum gate-switching energy of the most advanced 10-nanometer CMOS technology is projected to be about 8 x 10^{-16} J, operating up to about 10 GHz. More than half of this energy is devoted to powering interconnects between transistors. Tunneling field-effect transistors (TFETs) are another beyond-CMOS technology being pursued for high-speed and low switching energy. Devices based on magnetic spin can operate with very low energy dissipation but only at lower speeds.

The ultralow switching energy of SFQ is only achieved at cryogenic temperatures of about 4 K. From a system perspective, the energy required for refrigeration needs to be taken into account. Modern-day, closed-cycle cryocoolers (e.g., SHI Cryogenics [14]) can readily support projected SFQ processor cooling needs. No ongoing supply of liquid helium is necessary. With efficiencies of roughly 1,000 watts of wall plug power to provide 4 K cooling with 1 watt of heat dissipation, the effective SFQ switching energy is about 10^{-16} J, which is still nearly an order of magnitude lower than state-of-the-art CMOS.

Superconductive integrated circuit fabrication

Just as with silicon CMOS technology, the ability to fabricate superconductive electronics as planar integrated circuits is crucial to realizing complex and miniature SFQ processors. In fact, many of the

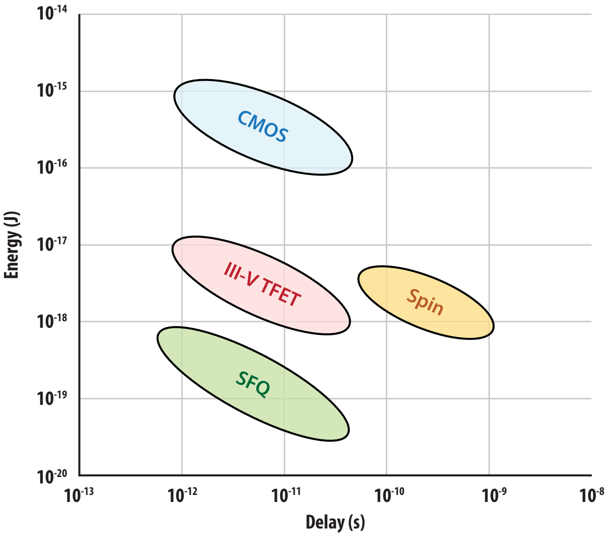


FIGURE 4. The projected gate-switching energy and delay time for several beyond-CMOS technologies (i.e., SFQ, III-V TFET, and Spin) compared with state-of-the-art silicon CMOS illustrates that SFQ switching energy is nearly an order of magnitude lower than state-of-the-art CMOS. [III-V refers to compounds with elements from both columns III and V of the periodic table.]

basic fabrication steps and processing tools for SFQ are borrowed directly from standard silicon integrated circuit technology.

Today, the most widely used JJ technology for very-large-scale integrated SFQ circuits is based on trilayers of niobium/aluminum/aluminum oxide/niobium. In a typical fabrication sequence, the niobium base electrode, aluminum layer, and niobium counter electrode are deposited by sputtering. The ultrathin (about 1 nanometer) aluminum oxide insulating tunnel barrier is formed by partial oxidation of the aluminum layer during this trilayer deposition process. The junction critical current density J_c is set primarily by the thickness of the tunnel barrier. Individual JJ circuit elements are patterned by photolithography and reactive ion etching. Accurate targeting of junction critical current I_c is realized by precise and reproducible control of photolithography and etch processes, along with high accuracy and uniformity of the starting trilayer J_c.

Multilayer SFQ fabrication processes are required to build complex circuits with dense interconnections between cells or gates. A 10-layer niobium process, with the cross-section shown in figure 5, is under development at Massachusetts Institute of Technology (MIT) Lincoln Laboratory. Circuits are fabricated on eight-inch silicon wafers with a single JJ layer. Superconductive niobium wiring layers are patterned for inductors. Metal wiring layers are separated by dielectric, and vias are used to interconnect layers to form circuits. A separate resistive layer is deposited and patterned for shunt resistors. All layers can be patterned by photolithography and etching. Chemical-mechanical planarization is employed at various steps in the process to maintain yield and uniformity. The lower metal layers can be used for interconnections between cells with passive transmission lines. A photomicrograph of a small portion of a fabricated SFQ chip fabricated at MIT Lincoln Laboratory in an eight-layer process is shown in figure 6.

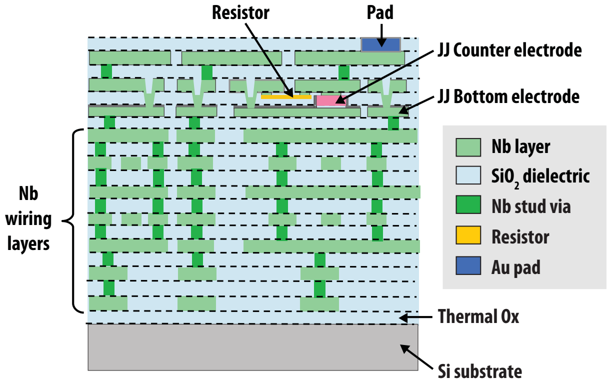


FIGURE 5. This diagram shows a cross section of a 10-layer niobium (Nb) SFQ microfabrication process under development at MIT Lincoln Laboratory. The circuits are fabricated on silicon (Si) substrates with a thermal oxide (Ox) layer. The Nb wiring layers are separated by silicon dioxide (SiO_2) dielectric layers. Gold (Au) pads are used for wirebond attachment to external circuits.

Several foundry processes have been developed worldwide and circuits with tens of thousands of JJs have been successfully demonstrated. In the US, Hypres, Inc. [15, 16] offers a four-layer standard process, with an option for six niobium layers, patterned with deep-ultraviolet photolithography to realize smaller feature sizes. The Superconductivity Research Laboratory of the International Superconductivity Technology Center (ISTEC) in Japan has developed a 10-layer niobium process (ADP2) for large-scale SFQ circuit demonstrations [17]. In Europe, the FLUXONICS Foundry for superconductive circuits has developed a three-layer niobium process [18]. A major emphasis of their work is on SFQ interface electronics for high-performance superconductor sensor applications.

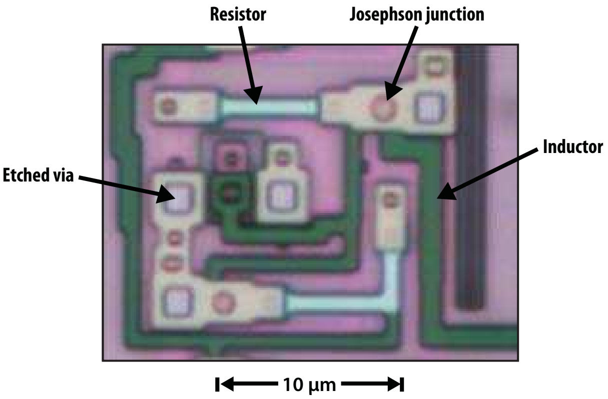


FIGURE 6. This photomicrograph shows an SFQ integrated circuit (at partial completion for visibility) with key circuit elements indicated.

Building blocks for SFQ processors

To date, the unique properties of SFQ electronics have already been exploited to demonstrate complex circuits for diverse applications. For radar and communications, a large effort in the US developed SFQ analog-to-digital converters based fundamentally on the speed and precision of flux quantization [19]. High-precision superconductor analog-to-digital converters comprised of approximately 5,000 JJs have been demonstrated at sampling rates of tens of gigasamples per second [20]. This work has been extended to successfully demonstrate [21] complete SFQ-based multichannel receivers including both digitizers and digital-signal processors on chips comprising about 11,000 JJs. Other efforts, particularly in Europe, have focused on SFQ implementations of readout circuits for cryogenic detectors [22]. Finally, SFQ circuits are being investigated as candidates for readout and control of superconductive qubits for quantum computing.

For high-performance computing applications, work has focused on developing computation building blocks such as adders and multipliers, which in turn can be configured into arithmetic logic units for general-purpose processors. Early efforts, initiated over a decade ago, included technology demonstrations of complete, but simple, microprocessors including the FLUX [23] and CORE [24] chips, with approximately 65,000 and 11,000 JJs, respectively. Efforts now are emphasizing scalable high-performance designs and architectures.

Significant work is under way on SFQ arithmetic circuits. A 16-bit sparse-tree adder, for example, was recently demonstrated by researchers from Stony Brook University, Yokohama National University, and Nagoya University [25]. The adder, designed to operate with rapid single-flux quantum (RSFQ) logic at 30 GHz with latency, the time required to calculate the output sum, of only 352 ps, is comprised of 9,941 JJs occupying an area of 8.5 square millimeters (mm^2). The same team has also demonstrated a low-latency eight-bit multiplier in the ISTEC 10 kA/cm^2 1.0 µm fabrication process [26]. The RSFQ multiplier operates at 20 GHz with a latency of only 447 ps and is comprised of 5,948 JJs in an area of 3.5 mm^2.

Path to ultralow power

Over the past several years, new SFQ design approaches have been demonstrated which promise to bring circuit power consumption down close to theoretical limits. These advances in SFQ technology are key to addressing the energy efficiency needs of high-performance computing.

While SFQ circuits have fundamentally low dynamic switching energy or power, the high static or standby power associated with providing dc bias currents in earlier SFQ circuits would typically dominate the power budget. A typical SFQ gate, such as the D-flip-flop described above, requires a bias current I_b through each junction during gate operation where I_b is typically comparable to I_c. Until recently, most SFQ circuit designs, including the SFQ logic family RSFQ, employed a resistive bias network to provide a stable dc bias current to each junction [8]. This resulted in a static bias power of about 800 nanowatts (nW) per gate which, for a gate operating as fast as 20 GHz, is about 60 times higher than the dynamic power dissipation [27]. For demonstrating the high-speed capability of SFQ circuits of modest complexity, this static power dissipation was not a major concern; however, minimizing energy usage is essential for large-scale SFQ computing applications.

New SFQ design families reduce the static power consumption to near zero. These techniques all involve removing or minimizing resistive circuit elements in the bias network. Some low-power design approaches, such as energy-efficient rapid single-flux quantum (ERSFQ), aim to reduce power with changes in bias circuitry only, while otherwise using existing RSFQ gate designs. Other recently developed approaches for ultralow power, such as reciprocal quantum logic (RQL), are based on entirely new logic designs.

In ERSFQ [28], and a similar approach called eSFQ [29], dc bias currents are delivered to SFQ gates via a superconductive bias network where resistors are replaced by current-limiting JJs feeding each gate. As the JJ current-voltage (*I*-*V*) curve in figure 1 indicates, with a very small dc applied voltage, the dc current will remain very close to the critical current I_c. Clocking schemes devised for ERSFQ and eSFQ ensure that the correct current-limiting operation of the bias JJs is maintained during dynamic gate-switching operation. In addition, bias line inductors are employed in these schemes to minimize bias current fluctuations during circuit operation. The size of these inductors, particularly for ERSFQ, can be relatively large, thus impacting overall logic density.

Another ultralow power SFQ design approach is RQL [30]. Here, dc bias lines are replaced by multiphase alternating-current (ac) power lines. All gates are inductively coupled to the ac power line and therefore no static power is dissipated on chip. The multiphase ac bias also serves as the clock for RQL circuits, replacing SFQ-base clock distribution networks.

An ultralow power RQL adder has been demonstrated by Northrop Grumman [31]. The eight-bit design employs a Kogge-Stone carry-look-ahead architecture and is implemented on a 5 mm x 5 mm chip with 815 junctions. The circuit was fabricated in the Hypres foundry process. Operating at a clock rate of 6.2 GHz, the power dissipation was only 510 nW. The authors project that a 64-bit adder would have a latency of only two clock cycles at 20 GHz in a more advanced fabrication process.

Path to higher integration

To meet future supercomputer needs, any beyond-CMOS technology must be scalable to high levels of integration. Reported exascale hardware designs envision SFQ processors with 20 million JJs [11]. This calls for circuit densities of at least one million JJs per square centimeter, nearly 10 times higher than integration levels typically reported today. High SFQ circuit density can be achieved through combining advances in several different aspects of SFQ technology as indicated in the development progression shown in figure 7.

As a first step, providing a sufficient number of metal wiring layers to enable efficient circuit layouts is essential for achieving high density. The 10-layer process shown in figure 5, for example, allows vertical stacking of circuit interconnections under logic cells, resulting in significant savings in chip area.

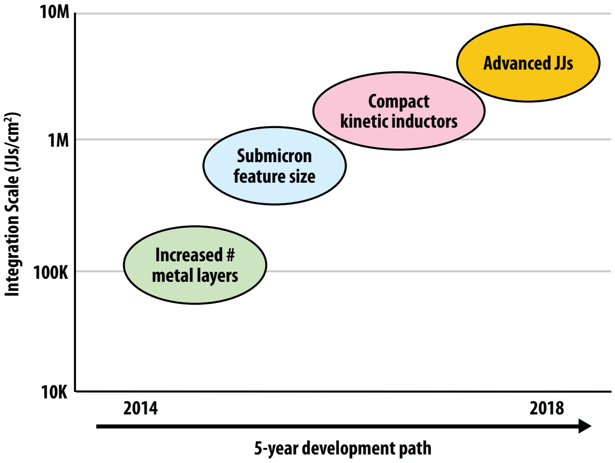


FIGURE 7. This diagram shows the development path for highly integrated SFQ processors.

With this approach, compact vias are also important for layer interconnections.

Reducing feature size is also an important step; however, the scaling rules for SFQ integrated circuits are very different than those for CMOS-based transistor circuits. Looking back at the photomicrograph in figure 6, the relative sizing of the circuit elements of a typical SFQ cell are readily apparent. The JJs, which are approximately 1 μm in diameter, are much smaller than the inductor and shunt resistor. While increasing junction J_c's and shrinking junction diameter could increase speed and therefore computational throughput, only a modest improvement in circuit density would result. This is because inductance values vary inversely with I_c, and I_c is generally constrained by noise and energy considerations.

Fortunately, reducing feature size is a way to shrink inductors. A typical SFQ cell inductor with a 1 μm-wide metal trace would require inductor lengths as long as approximately 10 μm. By reducing feature sizes down to 0.5 μm, the same inductance values could be reached in less than half the area.

Another way to shrink inductors is to take advantage of the so-called kinetic inductance of superconductors. This is a result of the physical momentum related to ballistic transport of charge in superconductors, which has no counterpart in normal conductors. Compact kinetic inductors fabricated from the superconductor niobium nitride look particularly attractive [32].

Finally, there are two promising research directions in advanced JJs which could provide a further boost to circuit density. The first is on self-shunted junctions which would eliminate the need for external shunt resistors. Forming JJs using amorphous niobium-silicon barrier layers is one approach being pursued [33]. Another research direction exploits the properties of JJs with a ferromagnetic tunnel barrier [34, 35]. This special type of device, called a π-junction, has a built-in π phase shift of the superconductive wave function. This reduces the phase shift necessary to generate across the inductor during gate switching, thus reducing the necessary inductance value. These ferromagnetic junctions are also being pursued as an SFQ-compatible memory element.

Conclusion

The past several years have seen major strides in the development of advanced superconductive SFQ digital electronics to meet the needs of future energy-efficient, high-performance computer systems. Fast SFQ computational circuits have been demonstrated operating on 100 times less power than comparable silicon CMOS circuits. New fabrication processes and circuit designs promise increased circuit complexity and integration.

About the author

The Massachusetts Institute of Technology Lincoln Laboratory is a federally funded research and development center that applies advanced technology to problems of national security. Research and development activities focus on long-term technology development as well as rapid system prototyping and demonstration. These efforts are aligned within its key mission areas: space control; air and missile defense technology; communication systems; cybersecurity and information sciences; intelligence, surveillance, and reconnaissance systems and technology; advanced technology; tactical systems; homeland protection; air traffic control; and engineering. The laboratory works with industry to transition new concepts and technology for system development and deployment.

References

[1] NVIDIA. “NVIDIA’s next generation CUDA compute architecture: Kepler GK110” [white paper]. 2012. Available at: http://www.nvidia.com/content/PDF/kepler/NVIDIA-Kepler-GK110-Architecture-Whitepaper.pdf.

[2] Borkar S, Chien A. “The future of microprocessors.” *Communications of the ACM.* 2011;54(5):67–77. doi: 10.1145/1941487.1941507.

[3] Bernstein K, Cavin RIII, Porod W, Seabaugh A, Welser J. “Device and architecture outlook for beyond CMOS switches.” *Proceedings of the IEEE.* 2010;98(12):2169–2184. doi: 10.1109/JPROC.2010.2066530.

[4] Cavin RIII, Lugli P, Zhirnov V. “Science and engineering beyond Moore’s law.” *Proceedings of the IEEE.* 2012;100:1720–1749. doi: 10.1109/JPROC.2012.2190155.

[5] Nikonov D, Young I. “Overview of beyond-CMOS devices and uniform methodology for their benchmarking.” *Proceedings of the IEEE.* 2013;99. doi: 10.1109/JPROC.2013.2252317.

[6] TOP500.org. “China’s Tianhe-2 supercomputer takes no. 1 ranking on 41st TOP500 list” [press release]. 2013. Available at: http://top500.org/blog/lists/2013/06/press-release/.

[7] Likharev K, Semenov V. “RSFQ logic/memory family: A new Josephson-junction digital technology for subterahertz-clock-frequency digital systems.” *IEEE Transactions on Applied Superconductivity.* 1991;1(1):3–28. doi: 10.1109/77.80745.

[8] Bunyk P, Likharev K, Zinoviev D. “RSFQ technology: Physics and devices.” *International Journal of High Speed Electronics and Systems.* 2001;(11)1:257–305. doi: 10.1142/S012915640100085X.

[9] Likharev K. “Superconductor digital electronics.” *Physica C.* 2012;482: 6–18. doi: 10.1016/j.physc.2012.05.016.

[10] Chen W, Rylyakov A, Patel V, Lukens J, Likharev K. “Rapid single flux quantum T-flip flop operating up to 770 GHz.” *IEEE Transactions on Applied Superconductivity.* 1999;9(2):3212–3215. doi: 10.1109/77.783712.

[11] Holmes S, Ripple A, Manheimer M. “Energy-efficient superconducting computing—power budgets and requirements.” *IEEE Transactions on Applied Superconductivity.* 2013;23(3). doi: 10.1109/TASC.2013.2244634.

[12] Van Duzer T, Turner CW. *Principles of Superconductive Devices and Circuits,* 2nd Edition. Upper Saddle River (NJ): Prentice Hall; 1999. ISBN-13: 978-0132627429.

[13] Bardeen J, Cooper LN, Schrieffer JR. "Microscopic theory of superconductivity." *Physical Review.* 1957;106(1):162–164. doi: 10.1103/PhysRev.106.162.

[14] SHI Cryogenics Group. Available at: http://www.shicryogenics.com/.

[15] Hypres, Inc. "Niobium integrated circuit fabrication, process #S45/200, design rules." Available at: http://www.hypres.com/foundry/niobium-process/.

[16] Hypres, Inc. "Niobium integrated circuit fabrication, process #03-10-45, design rules." Available at: http://www.hypres.com/foundry/niobium-process/.

[17] Nagasawa S, Satoh T, Hinode K, Kitagawa Y, Hidaka M, Akaike H, Fujimaki A, Takagi N, Yoshikawa N. "New Nb multi-layer process for large-scale SFQ circuits." *Physica C.* 2009;469:1578–1584. doi: 10.1016/j.physc.2009.05.219.

[18] Junert J, Brandel O, Linzen S, Wetzstein O, Toepfer H, Ortlepp T, Meyer H-G. "Recent developments in superconductor digital electronics technology at FLUXONICS foundry." *IEEE Transactions on Applied Superconductivity.* 2013;23(5):2013. doi: 10.1109/TASC.2013.2265496.

[19] Mukhanov O, Gupta D, Kadin A, Semenov V. "Superconductor analog-to-digital converters." Proceedings *of the IEEE.* 2004;92(10):1564–1584. doi: 10.1109/JPROC.2004.833660.

[20] Vernik IV, Kirichenko D, Filipov T, Talalaevskii A, Sahu A, Inamdar A, Kirichenko A, Gupta D, Mukhanov O. "Superconducting high-resolution low-pass analog-to-digital converters." *IEEE Transactions on Applied Superconductivity.* 2007;17(2):442–445. doi: 10.1109/TASC.2007.898613.

[21] Mukhanov O, Kirichenko D, Vernik I, Filipov T, Kirichenko A, Webber R, Dotsenko V, Talalaevskii A, Tang J, Sahu A, Shevchenko P, Miller R, Kaplan S, Sarwana S, Gupta D. "Superconductor digital-RF receiver systems." *IEICE Transactions on Electronics.* 2008;E91-C(3):306–317. doi: 10.1093/ietele/e91-c.3.306.

[22] Kunert J. "European roadmap on superconductive electronics—status and perspectives." *Physica C.* 2010;470:2079–2126. doi: 10.1016/j.physc.2010.07.005.

[23] Bunyk P, Leung M, Spargo J, Dorojevets M. "FLUX-1 RSFQ microprocessor: Physical design and test results." *IEEE Transactions on Applied Superconductivity.* 2003;13(2):433–436. doi: 10.1109/TASC.2003.813890.

[24] Yamanishi Y, Tanak M, Akimoto A, Park H, Kamiya Y, Irie N, Yoshikawa N, Fujimaki A, Terai H, Hashimoto Y. "Design and implementation of a pipelined bit-serial SFQ microprocessor, CORE1 β." *IEEE Transactions on* Applied *Superconductivity.* 2007;17(2):474–477. doi: 10.1109/TASC.2007.898606.

[25] Dorojevets M, Ayala C, Yoshikawa N, Fujimaki A. "16-bit wave-pipelined sparse-tree RSFQ adder." *IEEE* Transactions *on Applied Superconductivity.* 2013;23(3). doi: 10.1109/TASC.2012.2233846.

[26] Dorojevets M, Kasperek A, Yoshikawa N, Fujimaki A. "20-GHz 8 x 8-bit parallel carry-save pipelined RSFQ multiplier." *IEEE Transactions on Applied Superconductivity.* 2013;23(3). doi: 10.1109/TASC.2012.2227648.

[27] Mukhanov O. "Energy-efficient single flux quantum technology." *IEEE Transactions on Applied* Superconductivity. 2011;21(3):760–769. doi: 10.1109/TASC.2010.2096792.

[28] Kirichenko D, Sarwana S, Kirichenko A. "Zero static power dissipation biasing of RSFQ circuits." *IEEE* Transactions *on Applied Superconductivity.* 2011;21(3):776–779. doi: 10.1109/TASC.2010.2098432.

[29] Volkmann M, Sahu A, Fourie C, Mukhanov O. "Implementation of energy efficient single flux quantum digital circuits with sub-aJ/bit operation." *Superconductor Science and Technology.* 2013;26. doi: 10.1088/0953-2048/26/1/015002.

[30] Herr Q, Herr A, Oberg O, Ioannidis A. "Ultra-low-power superconductor logic." *Journal of Applied Physics.* 2011;109(10). doi: 10.1063/1.3585849.

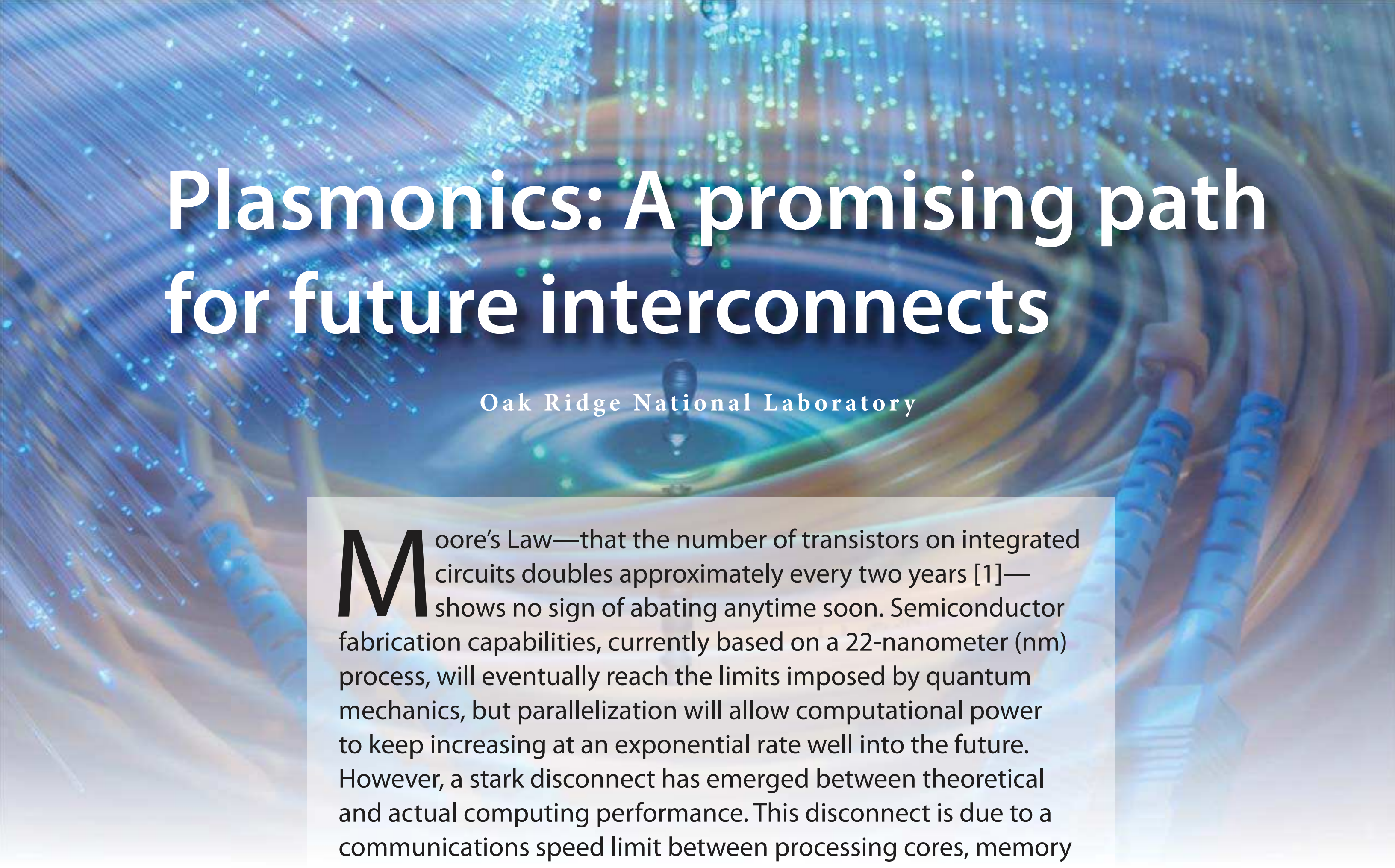
[31] Herr A, Herr Q, Oberg O, Naaman O, Przybysz J, Borodulin P, Shauck S. "An 8-bit carry look-ahead adder with 150 ps latency and sub-microwatt power dissipation at 10 GHz." *Journal of Applied Physics.* 2013;113. doi: 10.1063/1.4776713.

[32] Annunziata A, Santavicca D, Frunzio L, Catelani G, Rooks M, Frydman A, Prober D. "Tunable superconducting nanoinductors." *Nanotechnology.* 2010;21. doi: 10.1088/0957-4484/21/44/445202.

[33] Olaya D, Dresselhaus P, Benz S, Herr A, Herr Q, Ioannidis A, Miller D, Kleinsasser A. "Digital circuits using self-shunted Nb/NbxSi1-x/Nb Josephson junctions." *Applied Physics Letters.* 2010;96(21). doi: 10.1063/1.3432065.

[34] Feofanov A, Oboznov V, Bol'ginov V, Lisenfeld J, Poletto S, Ryazanov V, Rossolenko A, Khabipov M, Balashov D, Zorin A, Dmitriev P, Koshelets V, Ustinov A. "Implementation of superconductor/ferromagnet/superconductor pi-shifters in superconducting digital and quantum circuits." *Nature Physics.* 2010;6:593–597. doi: 10.1038/nphys1700.

[35] Ryazanov V, Bol'ginov V, Sobanin D, Vernik I, Tolpygo S, Kadin A, Mukhanov O. "Magnetic Josephson junction technology for digital and memory applications." *Physics Procedia.* 2012;36:35–41. doi: 10.1016/j.phpro.2012.06.126.



Plasmonics: A promising path for future interconnects

Oak Ridge National Laboratory

Moore's Law—that the number of transistors on integrated circuits doubles approximately every two years [1]—shows no sign of abating anytime soon. Semiconductor fabrication capabilities, currently based on a 22-nanometer (nm) process, will eventually reach the limits imposed by quantum mechanics, but parallelization will allow computational power to keep increasing at an exponential rate well into the future. However, a stark disconnect has emerged between theoretical and actual computing performance. This disconnect is due to a communications speed limit between processing cores, memory cache, and storage, and it is the primary performance bottleneck that prevents harnessing available computational power.

The interconnects at the heart of today's microprocessors consist of copper wires measuring several tens of nanometers to a few millimeters in dimension depending on their function. One successful strategy that has provided performance gains over the past 40 years has been to increase the clock speed of the microprocessor. Clock speeds have increased by three orders of magnitude in 40 years but have begun to level off, at around 4–5 gigahertz, within the past five years. This is due to the success of another strategy which involves reducing the size of the transistors themselves. As more transistors are crammed onto a chip, the dimensions of the transistors and the copper wires that supply power, clock, and instruction signaling must also be reduced.

Unfortunately, as copper wires shrink in size and the signaling frequency increases, the wires exhibit both more electrical resistance and signal propagation delay. Chip designers have countered with lower signaling voltages and ingenious signaling network layouts, but the communications bottleneck will endure as long as copper wires remain the interconnect of choice.

By 2017 the electrical current required for chip power and signaling will likely exceed the material limits of copper metal according to the International Technology Roadmap for Semiconductors (ITRS) 2011 report [2]. Although numerous efforts to identify potential copper material replacements are under way, retaining the same interconnect strategy is likely to gain only an order of magnitude increase in performance, representing a literal kicking the can down the road. Technologies have been proposed to replace electrical signaling over copper, a full account of which can be found in the ITRS reports. Photonics is one of these proposed replacements because photons can be generated easily, detected efficiently, and can have bandwidths three orders of magnitude greater than achievable with electrical signaling.

Silicon photonics is often heralded as the next generation of interconnect technology, capable of relatively low-loss, high-bandwidth optical signaling using integrated waveguides with cross sections as small as 200 nm x 200 nm. This size, while relatively small, is still an order of magnitude larger than today's current semiconductor fabrication process. If photonics, or any other potential technology, is to be a realistic contender to replace copper interconnects, it must have comparable dimensions to the transistor. Photonics is constrained by the diffraction limit: No dielectric element, resonator, or waveguide can have dimension less than half of the wavelength and still faithfully confine photons and function properly. Can we harness the advantages of optics at the nano-scale?

Yes—with plasmonics, a multidisciplinary field combining optics and solid-state physics, the study of optical fields bound to metal surfaces. A plasmon is a quantum mechanical quasiparticle consisting of collective oscillations of a metal's conduction electrons. A surface plasmon polariton (SPP) is a surface-bound propagating plasmon mode. One of the most intriguing aspects of plasmonics is for devices that operate well below the optical diffraction limit. This is due to some of the electromagnetic field energy being coupled to the metal's electron kinetic energy instead of being stored in the magnetic field.

As a result, plasmonic devices can effectively "squeeze" light into metal structures tens of nanometers in size, exceeding the optical diffraction limit by factors of 10 or more, yet retaining the large optical bandwidth. This is what makes plasmonics a promising contender to replace the copper interconnects in current architectures. Future plasmonic interconnects will require 1) sources of plasmons, 2) low-loss and high-confinement waveguides, and 3) plasmon detectors. In this article, we will provide a brief introduction and review of these three important technologies that will provide the necessary foundation for future plasmonic interconnects.

Plasmon sources

SPP sources are divided into two main technologies: nanolasers, which generate photons that are then coupled to plasmons, and SPASER-type devices, which are true nanoscale sources of plasmons. (SPASER stands for surface plasmon amplification of stimulated emission of radiation.) The cavity size of traditional lasers has a fundamental minimum, typically on the order of a few microns, imposed by the optical diffraction limit and round-trip cavity losses. As such, even state-of-the-art lasers are unsuitable for integration into chip architectures due to their size—they are simply too large.

Nanolasers exhibit all of the desirable properties of their larger cousins but with one major difference: The addition of metal within the cavity allows the cavity to be much smaller (see figure 1). By tuning the laser emission close to an SPP resonance, the nanolaser experiences both a large increase in the effective modal gain and group refractive index. This allows stimulated emission within a much smaller cavity. Nanolasers have recently been demonstrated with a mode volume of 0.4 cubic wavelengths (λ^3) and operating under electrical injection at room temperature in the important telecom band [3]. Over the next two years, we are likely to see even more impressive demonstrations of nanolasing with devices having smaller mode volumes, a selection of wavelengths, and tunable cavity configurations. Within five years, we will likely witness demonstrations of pulsed nanolasers with subfemtosecond pulse widths and the beginnings of integration with advanced silicon photonic devices.

Nanolasers, with their subdiffraction-limit mode volumes, are not ideal SPP sources for two specific reasons: at least one dimension has to be greater than $\lambda/2$, and nanolasers emit photons instead of plasmons and thus will experience some photon-plasmon coupling loss. SPASERs, the surface plasmon analog to lasers, may very well be the ideal localized plasmon source capable of ultrafast attosecond operation [4, 5] but are a much less mature technology in comparison to nanolasers. The most widely studied SPASER devices are comprised of

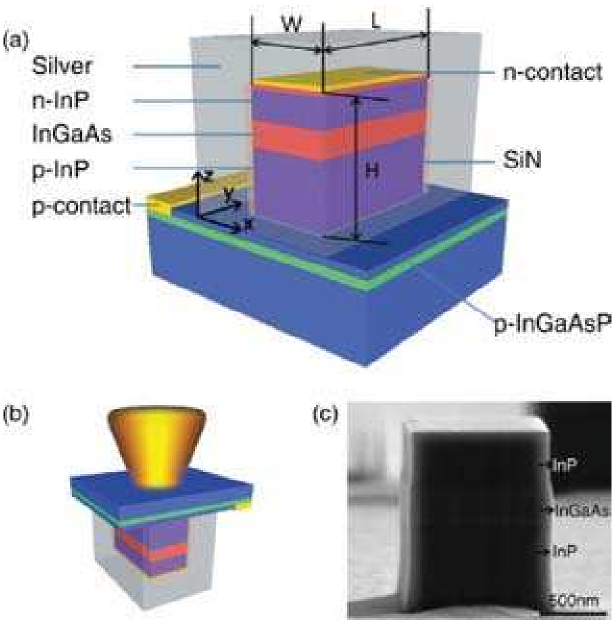


FIGURE 1. Nanolasers exhibit all of the desirable properties of their larger cousins but with one major difference: The addition of metal within the cavity allows the cavity to be much smaller. This schematic of a nanolaser (a) with light emission from the reverse side (b) consists of a semiconductor pillar (c) encapsulated with silver to form a metallic cavity. The semiconductor materials are as follows: indium gallium arsenide (InGaAs), silicon nitride (SiN), indium phosphide (InP), indium gallium arsenide phosphide (InGaAsP). Where the prefixes 'p-' and 'n-' denote positive and negative donor doping respectively, the metals are as follows: gold, platinum, and titanium. Reprinted figure with permission from Ding K, Liu ZC, Yin LJ, Hill MT, Marell MJH, van Veldhoven PJ, Nöetzel R, Ning CZ, *Physical Review B,* 85, 041301(R), fig.1, 2012. Copyright 2012 by the American Physical Society. Available at http://dx.doi.org/10.1103/PhysRevB.85.041301 [3]. Readers may view, browse, and/or download material for temporary copying purposes only, provided these uses are for noncommercial personal purposes. Except as provided by law, this material may not be further reproduced, distributed, transmitted, modified, adapted, performed, displayed, published, or sold in whole or part, without prior written permission from the American Physical Society.

metal nanoparticles, often 20–40 nm in diameter but as small as 1 nm, that are coated with a gain material (see figure 2). When excited, the gain medium supplies the energy for plasmon emission, with the nanoparticle itself acting as the resonant cavity. This results in the stimulated emission of coherent plasmons from the nanoparticle, albeit with a poorly defined emission profile. Placing a plasmonic waveguide close to a SPASER causes the plasmons generated in the SPASER to couple to the waveguide.

However, SPASERs have only been demonstrated using optical excitation with external pump lasers and not the electrical excitation so crucial to realize integrated devices. It turns out that electrically excited SPASERs are a rather difficult problem to solve due to the Purcell effect: The spontaneous emission rate, characterized by broadband and incoherent plasmon emission, increases as the resonator's cavity volume decreases. In order to counter the Purcell effect, more energy must be supplied to the gain medium to achieve stimulated plasmon emission (i.e., spasing). For resonator volumes as small as metal nanoparticles, the electrical current required for spasing is greater than the gain material can physically handle. Electrical spasing requires development of new gain materials, artificial materials with engineered metal-like properties, or novel device designs that somehow circumvent the Purcell effect.

Some workers in the plasmonics community argue that spontaneous emission, generally regarded as an unwanted artifact for SPASERs, might be useful as a plasmon source in itself. Surface plasmon-emitting devices (SPEDs) are sources of broadband and incoherent plasmons generated from spontaneous emission and are analogous to their optical cousins. While a number of researchers strive to develop the SPASER, it is likely that SPEDs would be of great value for certain applications where broadband yet incoherent plasmon output is beneficial. Such an application might work for a plasmonic interconnect scheme.

SPASERs and SPEDs represent ideal plasmon sources. They are truly nanoscale devices made of compatible materials and have comparable dimensions with the smallest microprocessor structure. We are likely to see significant SPASER advances in the next five years. The coupling efficiency of nanoparticle plasmons from SPASERs and SPEDs to SPP waveguides will improve considerably and manufacturing processes that allow reliable device fabrication will be developed. The most important development will be the demonstration of electrical SPASERs and will be foretold by developments in high-gain semiconductor materials. This is a necessity for nanoparticle SPASERs, yet we believe that the SPED may offer a worthwhile alternative for interconnect schemes requiring a broadband and power-efficient plasmon source.

Waveguides

As SPPs propagate on the surface of metals, a natural choice of an SPP waveguide would consist of a thin (< 50 nm) and narrow (< 200 nm) metal strip supported on a dielectric material. SPP waveguides are functional elements required to build more advanced devices such as splitters, couplers, modulators, and routers necessary for interconnect applications. In general, SPP waveguides support two types of SPP excitation: weakly bound long-range SPP modes (LRSPPs) and strongly bound short-range SPP modes (SRSPPs) regardless of the waveguide geometry.

There is one overarching generality with SPP waveguides no matter what the geometry—with increased plasmon confinement comes increased loss due to the metal itself. When an application requires very high confinement for ultradense signal propagation (e.g., intrachip core-core communication), one must deal with the higher losses imposed by a more-confined SRSPP mode. Conversely, when an application requires longer range propagation (e.g., intraboard chip-memory communication) without dense signaling or more crosstalk immunity, one can opt for the less-confined LRSPP mode.

The loss experienced by SPP modes can be reduced, potentially compensated for (zero-loss), and even amplified by using a gain medium in conjunction with a metal waveguide. Partial loss compensation has been demonstrated by a number of authors on a variety of SPP structures using optical excitation [6, 7] with external pump lasers (see figure 3). For total loss compensation, the gain supplied to the system must equal the plasmonic losses, which in turn depends on the device geometry and the SPP modes. Researchers estimate 1,000–100,000 cm^{-1} of gain is required for total loss compensation for SPP waveguides. In comparison, the highest gain achievable in semiconductor material is approximately 3,000 cm^{-1}, barely enough to provide full loss compensation for even LRSPPs.

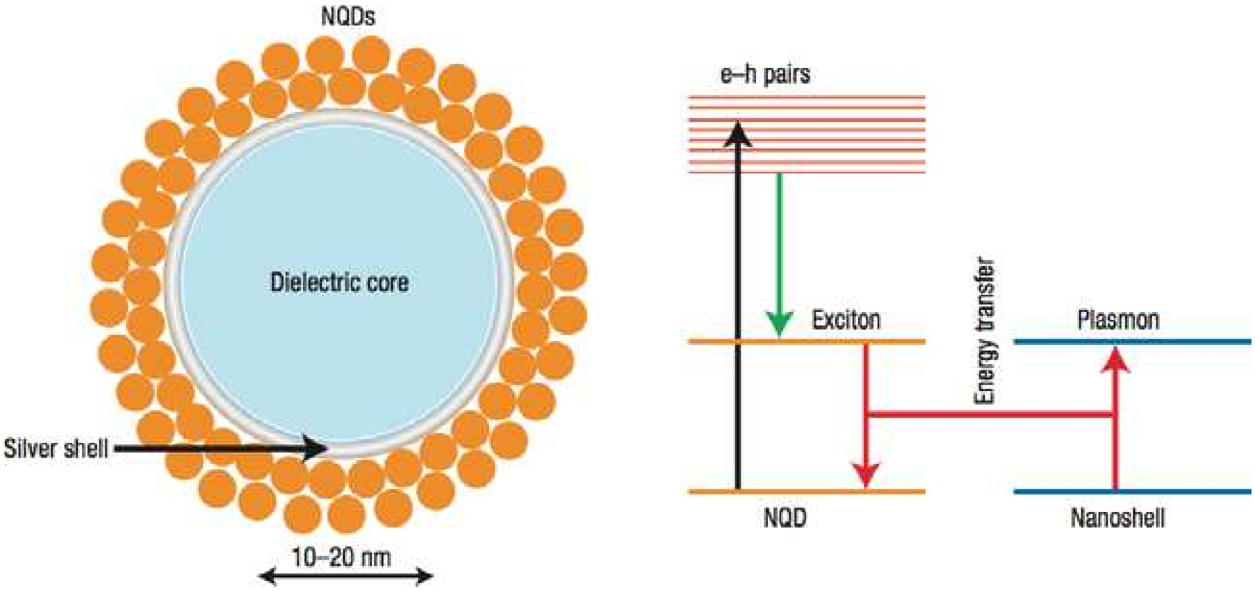


FIGURE 2. The most widely studied SPASER devices are comprised of metal nanoparticles that are coated with a gain material. This SPASER consists of a silver nanoshell particle surrounding a dielectric core and coated with nanocrystal quantum dot (NQD) gain medium (left). The spasing mechanism shows energy transfer between the gain medium and plasmon (right). Reprinted by permission from Macmillan Publishers Ltd: *Nature Photonics,* available at http://www.nature.com/nphoton/index.html, Stockman MI, "Spasers explained," doi: 10.1038/nphoton.2008.85, fig. 1 (a, b), 2008 [5].

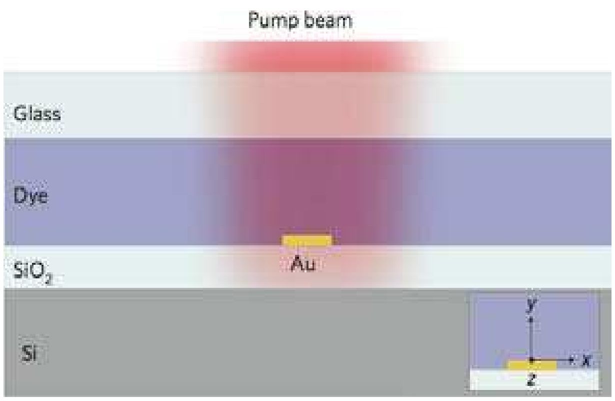


FIGURE 3. This cross section of an SPP waveguide device illustrates fractional loss compensation using optical excitation. Light from the pump beam (external to the device) excites dye molecules sandwiched between the silicon dioxide (SiO_2) fused quartz and glass layers—all supported on a silicon (Si) substrate. Plasmons propagating in the gold (Au) stripe waveguide are partially amplified (i.e., experience reduced losses) because of the excited dye molecules. As a result, the plasmon can propagate further thanks to the loss compensation. Reprinted by permission from Macmillan Publishers Ltd: *Nature Photonics,* available at http://www.nature.com/nphoton/index.html, De Leon I, Berini P, "Amplification of long-range surface plasmons by a dipolar gain medium," doi: 10.1038/nphoton.2010.37, fig. 1a, 2010 [7].

Given the integration of existing gain materials with SPP waveguides, we expect to see the first reports of loss compensated in LRSPP-mode waveguides using optical excitation within two years. This will require minimization of existing LRSPP-mode losses and a device design to maximize the plasmon field overlap with the gain medium. We anticipate this demonstration will spur development of new gain materials specifically for plasmonics that provide yet more gain, albeit in a narrow spectral range.

During the same period, we will see demonstrations of electrical excitation by current injection into the gain medium. This will be much more difficult in comparison to the optical excitation method. To prevent the available gain from being distributed over unwanted SPP modes, the waveguide has to be made small enough to support only the desired LRSPP or SRSPP mode(s), reducing the device volume, thus bringing the Purcell effect into play. Within five years, we anticipate electrical excitation of SPP waveguides that achieve total loss compensation for both LRSPP and SRSPP modes. We also anticipate the development of artificial metal-like materials that exhibit low loss over a very narrow wavelength range. These developments will in turn enable all other elements critical to a plasmonic interconnect scheme and will herald maturation of plasmonic technology.

Detectors

The third and final major component in any future plasmonic interconnect scheme is the plasmon detector, namely a device that signals the arrival of SPPs by the output of a macroscopic voltage or current pulse. Two promising detection schemes exist for plasmon detection: the Schottky barrier detector (SBD) [8] and the superconducting nanowire single-plasmon detector (SNSPlD) [9, 10]. Both detectors operate on the same principle by using the heat generated by decaying SPPs.

At the heart of an SBD is a semiconductor-metal junction, called a Schottky junction, which has an energy barrier height determined by the choice of semiconductor and metal (see figure 4). For example, gold on *n*-doped silicon has a Schottky barrier height of 0.83 electron volts. When a plasmon with energy greater than the Schottky barrier height decays, it leads to "hot carriers" being generated. These hot carriers overcome the Schottky barrier, generating an electrical current, which is detected by electronics and results in a plasmon detection event.

The sensitivity of SBDs depend on how small the Schottky junction can be made; large-area junctions suffer from high dark currents (i.e., false detection events) and are therefore less sensitive. Sensitivity can be increased by reducing the detector area, which in turn requires more plasmon decay in this smaller area. Therefore, sensitive SBDs can be realized using waveguides specifically designed to be extremely lossy (i.e., using the SRSPP mode). Other properties of SBDs, namely detection bandwidth and responsivity, can be tuned by using different metal-semiconductor combinations to alter the Schottky barrier height.

Current SBDs are capable of room-temperature operation and detection of plasmon power down to

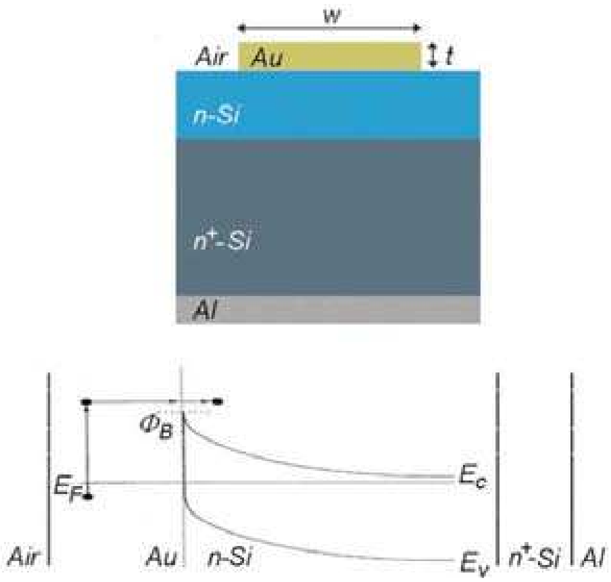


FIGURE 4. This cross section of a Schottky barrier plasmon detector (top) shows the gold (Au) waveguide and Au *n*-silicon (*n*-Si) Schottky barrier on an aluminum (Al) substrate. The junction energy level diagram (bottom) shows the Schottky barrier height (Φ_B). E_F, E_c, and E_v refer to the Fermi, conduction band, and valance band energies respectively, which are all dependent on the materials used. Reprinted with permission from *Applied Physics Letters,* Akbari A, Berini P, "Schottky contact surface-plasmon detector integrated with an asymmetric metal stripe waveguide," fig. 1, doi: 10.1063/1.3171937. Copyright 2009, AIP Publishing LLC [11].

−46 decibels per milliwatt (dBm) [12]. We expect device improvements to further reduce the dark current and increase sensitivity, with −60 dBm at a wavelength of 1.55 micrometers achievable within the next two years. For ultrasensitive and very low-power plasmon detection, even down to the few- or single-plasmon level, SBDs have a long development path in front of them. SBDs are intrinsically fast devices, capable of detecting gigahertz rates thanks to their heritage as sub-bandgap photodetectors, and do not exhibit a dead-time effect. Over the next five years, we expect SBDs to become more sensitive and responsive; in fact, we expect to see demonstrations of devices as sensitive as Geiger-mode SBDs. We anticipate SBDs to become and remain the dominant technology in plasmon detection in the future.

SBDs are unlikely, in the near future at least, to detect down to the few- and single-plasmon level. Such a capability would be extremely useful in applications where excessive loss is present, where loss-compensation schemes are unavailable, and also in emerging fields such as quantum plasmonics. For ultralow- to single-plasmon detection, we turn to the SNSPlD, itself also a derivative device from the photon-detection regime. SNSPlDs consist of a niobium-nitride nanowire structure grown in contact with or in close proximity to a plasmon waveguide. The entire device is cooled down to less than 10 kelvins (K; −263 °C), below the niobium-nitride superconducting critical temperature (T_c), using a liquid helium refrigerator. In an SNSPlD, the decaying plasmon locally heats the niobium-nitride nanowire, which results in a hotspot—a localized nonsuperconducting region with finite electrical resistance—and leads to a macroscopic current pulse.

SNSPlDs are extremely sensitive devices and, like their photon-detecting cousins, can detect single plasmons with efficiencies greater than 50% in the telecom wavelength band. The detector jitter, or timing uncertainty, is often very low at around 100 picoseconds, and SNSPlDs can operate at relatively high count rates exceeding hundreds of megahertz.

However, while all these detector specifications are impressive for single-plasmon counting applications, for high-bandwidth operation within a plasmonic interconnect scheme, one specific advance must be made for SNSPlDs to be a viable detector technology. The sub-10 K cooling is the most obvious drawback for SNSPlD detectors, and the only practical way for SNSPlDs to be used in interconnect technology is for new high-T_c superconductors to be developed. This is not an easy problem to solve, and while we anticipate current research on high-T_c superconductivity to keep bearing fruit, we do not expect a disruptive high-T_c superconductor breakthrough within the next five years. This unfortunately leaves SNSPlDs as a research tool primarily for single-plasmon applications and possibly interconnect schemes where the detectors can be located well away from the microprocessor.

Summary

Plasmonics offers the benefits of high-bandwidth signaling with physical confinement down to the true nanoscale due to the metals on which plasmons are supported. Plasmonic interconnects is a major contender as a future technology to alleviate the current communications bottleneck in computer architectures. We have introduced three major plasmonic technologies necessary to realize this vision: 1) plasmon sources such as nanolasers, SPASERs, and SPEDs; 2) low- and zero-loss plasmon waveguides, which in turn will be necessary for splitters, modulators, and routers; and 3) efficient plasmon detectors. However, the metal which permits nanoscale confinement of an optical field is also a great disadvantage due to losses incurred. Therefore, we anticipate some major breakthroughs within the next five years on development of gain materials and artificial metal-like materials specifically for plasmonics applications. This will result in demonstrations of low- and zero-loss waveguides and electrical injection spasing. We also expect developments in plasmonic detectors leading low-noise and high-sensitivity plasmon detection. These advances will likely propel plasmonics forward at a rapid pace, enabling practical plasmonic devices and bringing the plasmonic interconnect vision closer to reality.

About the author

This work was performed at **Oak Ridge National Laboratory (ORNL)**, operated by UT-Battelle, LLC for the US Department of Energy under contract number DE-AC05-00OR22725. Scientists and engineers at ORNL conduct basic and applied research and development to create scientific knowledge and technological solutions that strengthen the nation's leadership in key areas of science; increase the availability of clean, abundant energy; restore and protect the environment; and contribute to national security. Specific program thrusts include ultrascale computing, cyber/energy infrastructure protection, quantum information science, homeland protection, and difficult problems that face the intelligence community.

References

[1] Moore GE. "Cramming more components onto integrated circuits." *Electronics*. 1965;38(8):114–117. doi: 10.1109/jproc.1998.658762 [Reprint from *Proceedings of the IEEE*. 1998;86(1):82–85].

[2] ITRS. "International technology roadmap for semiconductors, 2011 edition." 2011. Available at: http://www.itrs.net/Links/2011ITRS/Home2011.htm.

[3] Ding K, Liu ZC, Yin LJ, Hill MT, Marell MJH, van Veldhoven PJ, Nöetzel R, Ning CZ. "Room-temperature continuous wave lasing in deep-subwavelength metallic cavities under electrical injection." *Physical Review B*. 2012;85(4). doi: 10.1103/PhysRevB.85.041301.

[4] Bergman DJ, Stockman MI. "Surface plasmon amplification by stimulated emission of radiation: Quantum generation of coherent surface plasmons in nanosystems." *Physical Review Letters*. 2003;90(2). doi: 10.1103/PhysRevLett.90.027402.

[5] Stockman MI. "Spasers explained." *Nature Photonics*. 2008;2:327–329. doi: 10.1038/nphoton.2008.85.

[6] Noginov MA, Podolskiy VA, Zhu G, Mayy M, Bahoura M, Adegoke JA, Ritzo BA, Reynolds K. "Compensation of loss in propagating surface plasmon polariton by gain in adjacent dielectric medium." *Optics Express*. 2008;16(2):1385–1392. doi: 10.1364/OE.16.001385.

[7] De Leon I, Berini P. "Amplification of long-range surface plasmons by a dipolar gain medium." *Nature Photonics*. 2010;4:382–387. doi: 10.1038/nphoton.2010.37.

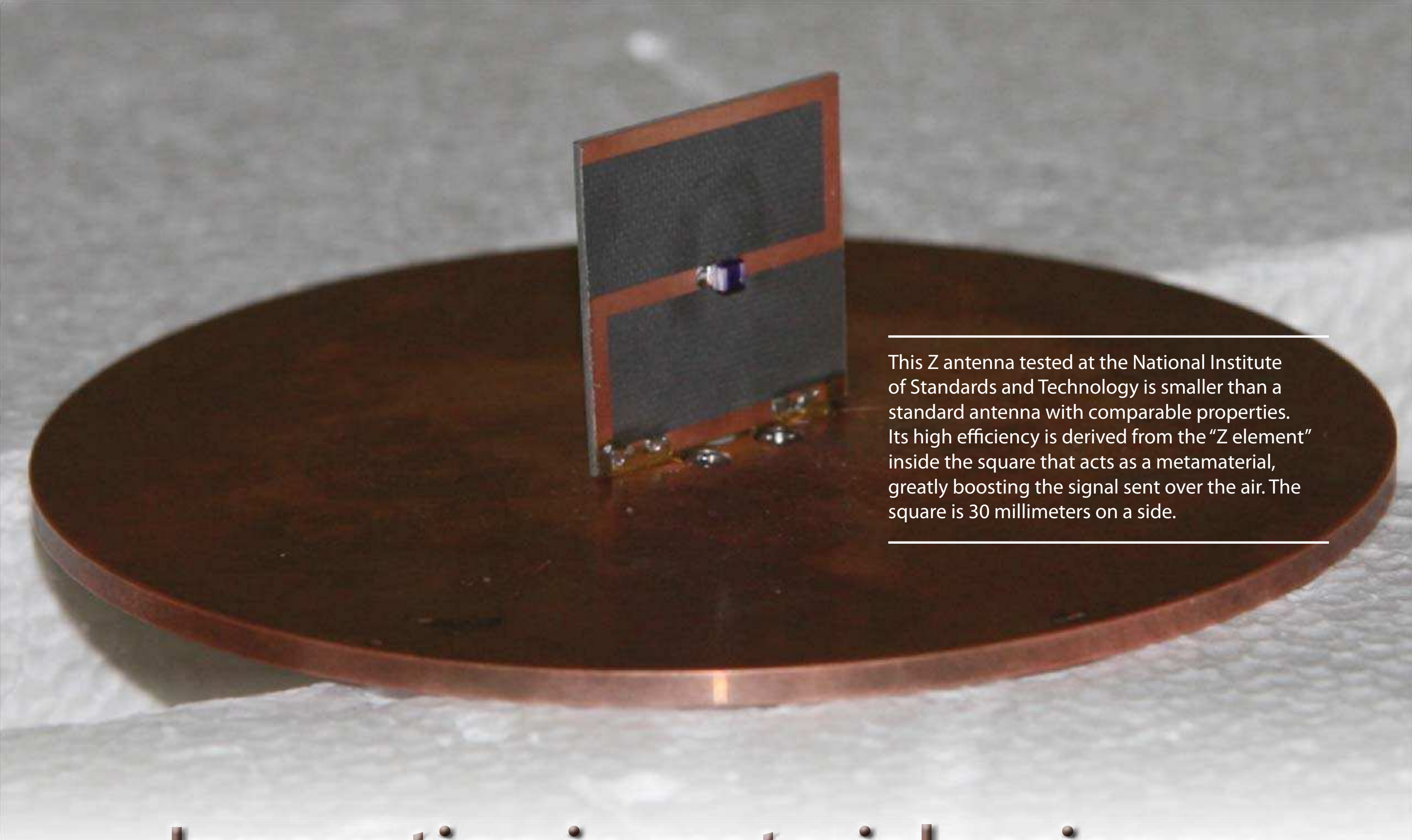
[8] Akbari A, Berini P. "Schottky contact surface-plasmon detector integrated with an asymmetric metal stripe waveguide." *Applied Physics Letters*. 2009;92(2). doi: 10.1063/1.3171937.

[9] Gol'tsman GN, Okunev O, Chulkova G, Lipatov A, Semenov A, Smirnov K, Voronov B, Dzardanov A, Williams C, Sobolewski R. "Picosecond superconducting single-photon detector." *Applied Physics Letters*. 2001;79(6):705–707. doi: 10.1063/1.1388868.

[10] Heeres RW, Dorenbos SN, Koene B, Solomon GS, Kouwenhoven LP, Zwiller V. "On-chip single plasmon detection." *Nano Letters*. 2010;10(2):661–664. doi: 10.1021/nl903761t.

[11] Akbari A, Berini P. "Schottky contact surface-plasmon detector integrated with an asymmetric metal stripe waveguide" (figure 1). *Applied Physics Letters*. 2009;95(2):021104. doi: 10.1063/1.3171937.

[12] Akbari A, Tait RN, Berini P. "Surface plasmon waveguide Schottky detector." *Optics Express*. 2010;18(8):8505–8514. doi: 10.1364/OE.18.008505.



This Z antenna tested at the National Institute of Standards and Technology is smaller than a standard antenna with comparable properties. Its high efficiency is derived from the "Z element" inside the square that acts as a metamaterial, greatly boosting the signal sent over the air. The square is 30 millimeters on a side.

Innovation in materials science: Electromagnetic metamaterials

Jane E. Heyes | Nathaniel K. Grady | Diego A. R. Dalvit | Antoinette J. Taylor

Material properties affect the propagation of electromagnetic (EM) waves in profound ways, which has allowed for devices ranging from eyeglasses to radar to fiber-optic cables. However, there are a limited number of responses found in natural materials. How can the range of possibilities be expanded? Enter EM metamaterials.

EM metamaterials are composites built from a structured combination of conductors, semiconductors, and insulators. The individual features make up an ordered array smaller than the wavelengths of radiation they are designed to affect, so the EM wave responds to the overall combination of these individual structures as if it were an effectively homogeneous material. By providing effective material properties not found in nature, metamaterials have the potential to aid in the creation of ultrathin planar lenses, superresolution microscopes, compact antennas, faster computer chips, and surfaces that radically alter or cloak the EM signature of an object (e.g., an invisibility cloak).

The limitations of natural materials are a major obstacle that must be overcome to meet the ever-increasing demand for faster, lighter, cheaper, and more compact devices, making metamaterials an important and timely tool for meeting future technology needs.

Background

Two fundamental EM properties of matter are the electric permittivity (ε) and permeability (μ). In all naturally occurring materials at all EM wavelengths, these two values are never simultaneously negative. Knowing the values of these two parameters, it is possible to calculate a number of different properties of the material, including the speed of propagation, propensity to absorb energy, ability to reflect, and possibly the effects on polarization.

The speed of propagation, inversely proportional to the refractive index, describes how a wave's path will change when it moves from one medium to another at an oblique angle. When a wave travels from a lower index medium into a higher one, it bends closer to the line normal to the interface between the two media, with the inverse true for a wave moving from a higher index medium into a lower one.

Over 40 years ago, Victor Veselago predicted that a negative index of refraction would result in light bending in the opposite direction from what is expected [1], but no natural materials have a negative refractive index. In 1999, John Pendry, a pioneer in metamaterials, worked on reducing the electrical plasma frequency in metal wires and created an artificial magnetic response via metallic split-ring resonators (SRRs), illustrated in figure 1, a key theoretical step in creating a negative refractive index [2].

David Smith and colleagues were the first to demonstrate composite metamaterials, using a combination of plasmonic-type metal wires and an SRR array to create a negative ε and negative μ in the microwave regime [3]. They demonstrated that EM waves (e.g., light) are able to propagate in such composite metamaterials with simultaneously negative effective values of the constitutive parameters ε and μ—that is, with a negative index of refraction.

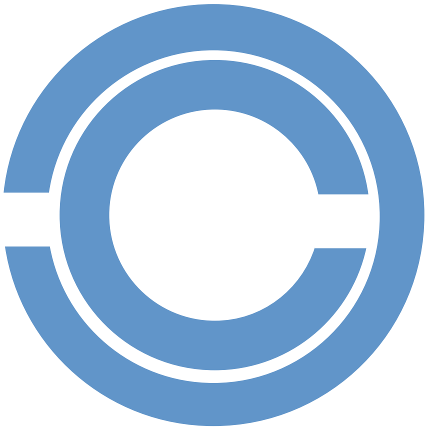


FIGURE 1. A schematic illustration of a split-ring resonator.

These are the kind of hypothetical "substances" that Veselago had speculated about in the past. In his paper, Veselago predicted several fundamental phenomena occurring in or in association with such substances, including the characteristic frequency dispersion, negative index of refraction, reversal of Snell's law, focusing with a flat slab, and reversal of Doppler effect and Cherenkov radiation—all of which have now been experimentally observed using metamaterials.

Thus, the study of metamaterials began with the exploration of materials with a negative refractive index. However, the bulk of research has diverged into different specialties, and now many different kinds of devices are studied over many decades of the EM spectrum. The expanded breadth of research has yielded discoveries of new phenomena including seminal proof-of-concept demonstrations of superresolution in optical imaging, perfect metamaterial absorbers, EM invisibility or cloaking, and transformation optics. Figure 2 shows examples of metamaterials [4].

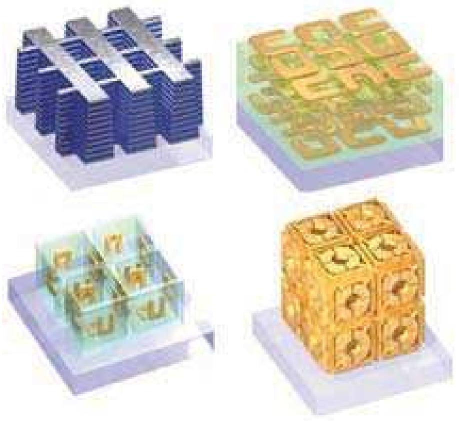


FIGURE 2. These example metamaterials are composites built from a structured combination of conductors, semiconductors, and insulators. Image adapted by permission from Macmillan Publishers Ltd: *Nature Photonics,* Soukoulis CM, Wegener M, "Past achievements and future challenges in the development of three-dimensional photonic metamaterials," doi: 10.1038/nphoton.2011.154, 2011 [4].

While the initial work primarily focused on three-dimensional metamaterials, recent efforts focusing on explicitly considering two-dimensional planar metamaterials (i.e., metasurfaces) and conceptually focusing on modifying boundary conditions at interfaces have proven extremely fruitful. For example, the generalized laws of refraction and the interference theory of perfect absorbers/antireflection coatings, both derived from investigating how the boundary conditions at the interface between two materials can be radically altered by metasurfaces, have led to rapid advances in the area of planar optics. Despite being a more recent innovation, metasurfaces are likely to rapidly reach significant commercial relevance due to their being readily fabricated using widespread conventional lithography techniques and being easier to integrate directly onto existing detectors or sources.

Metamaterial absorbers, emitters, and antireflection coatings

In 2007, researchers developed a metamaterial capable of absorbing all of the light that strikes it—a perfect absorber—representing one of the most important applications of metamaterials [5]. As a function of frequency ω, a material's effective impedance, defined as $Z(\omega) = [\mu(\omega)/\varepsilon(\omega)]^{1/2}$ changes. At a particular frequency, the impedance matches the free-space impedance (Z_0), and therefore reflection is minimized. In metamaterials with simultaneous electrical and magnetic resonances, both the effective permittivity ε(ω) and permeability μ(ω), are highly frequency dependent and can be tailored independently, making it much easier to achieve a high-reflection state. If the metamaterial also achieves high loss, resulting in low transmission, then near-unity absorption can occur.

Additional efforts in understanding the underlying physics responsible for the impedance matching and perfect absorption are also under way. Researchers at Los Alamos National Laboratory (LANL) recently composed an interference theory [6] (see figure 3c) and explained the observed perfect absorption and antiparallel surface currents in two metallic layers. This theoretical advance also led to the development of highly efficient ultrathin planar polarization rotators and brought the efficiency of structures exhibiting generalized refraction into the realm of practical devices, as discussed in more detail below.

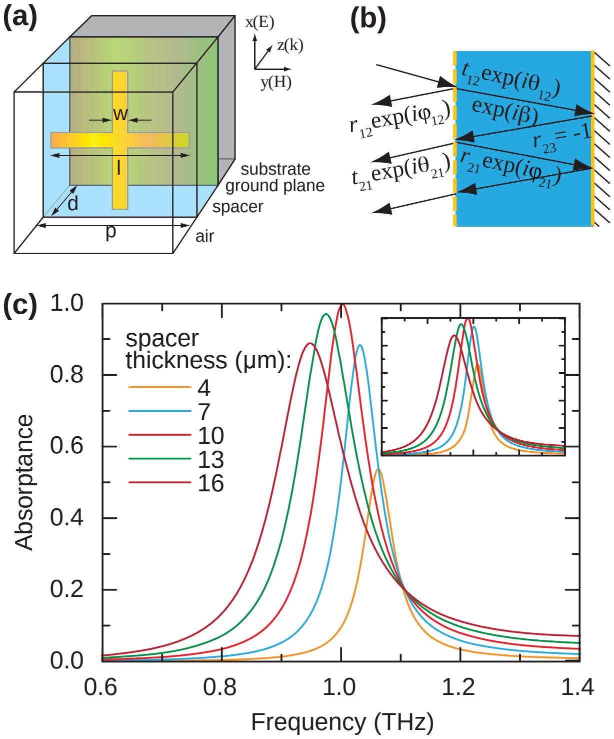


FIGURE 3. (a) This illustrated metamaterial perfect absorber consists of a metal cross-resonator array, dielectric spacer, metal ground plane, and substrate. **(b)** This diagram shows the interference model of metamaterial perfect absorption. **(c)** This graph shows absorptance in the decoupled metamaterial absorber using the interference model for various spacer thicknesses. The inset graph is a simulation of absorptance when treating the whole metamaterial absorber as a coupled system. Image adapted by permission from The Optical Society: *Optics Express,* Chen H, "Interference theory of metamaterial perfect absorbers," doi: 10.1364/OE.20.007165, 2012 [6].

Metamaterial perfect absorbers typically consist of a subwavelength resonator array backed with a metal ground plane and are separated with a dielectric spacer, as illustrated in figure 3(a). Compared to conventional absorption screens, the overall thickness of a metamaterial absorber is much smaller than the operation wavelength.

Currently, work in this field is focused on creating multiband and broadband metamaterial absorbers. These typically employ multilayered metamaterials or unit cells containing structures resonating at different frequencies.

Metamaterial-based absorbers are expected to increase energy conversion efficiency in photovoltaics and solar-thermal energy harvesting systems. The design of nanostructured "black" superabsorbers from materials comprised only of lossless dielectric materials and highly reflective noble metals represents a new research direction. For example, Harry Atwater at the California Institute of Technology is currently investigating metal–insulator–metal stack-based metamaterial absorbers with the intention of increasing the efficiency of photovoltaic or thermovoltaic cells [7].

Regarding the terahertz frequency range, there are also efforts to produce narrowband terahertz sources with relative high output power [8]. Liu et al. found that the emissivity (i.e., the ability of a material's surface to emit heat as radiation) and absorptivity of a surface follows Kirchhoff's law of thermal radiation of blackbody [9]. These results may have a great effect on controlling thermal signatures emitted from an object. For example, the outer surface of a hot object can be coated with a designed metamaterial to control the emissivity in a narrow spectral frequency which will deviate the natural thermal blackbody spectrum. This concept also has the potential to enable the creation of high-efficiency incandescent light sources [10] and play a key role in thermophotovoltaics.

Polarization control

The polarization state is one of the basic properties of EM waves conveying valuable information that is important in transmitting signals and making sensitive measurements. In fact, EM polarization has greatly affected our daily life for products as simple as sunglasses to high-tech applications including radar, laser technology, fiber-optic communications, liquid-crystal displays, and three-dimensional movies. Similar to controlling the EM wave intensity, manipulating the polarization states enables many applications, and its importance should not be underestimated.

As such, there has been a long history in the development of numerous devices for manipulation of EM polarization states, including polarizers, half-wave plates, and quarter-wave plates. Conventional approaches include using gratings, birefringent crystals, and Brewster plates. In order to code information into the polarization state, actively controllable wave plates are highly desired to modulate the beam polarization.

However, there are still many challenges in the development of polarimetric devices. Birefringent crystals work well for short wavelengths in the visible and near-infrared regimes but become bulkier and costlier to fabricate or have other undesirable properties, such as being soft or hydroscopic, for longer wavelengths. They also suffer from a narrow operating bandwidth, a problem sometimes ameliorated with complex fabrication. Polymers are also widely used in polarimetric applications; however, they are not suitable for longer wavelengths due to high absorption.

Scientists in Antoinette Taylor's group at LANL carried out research on improving the efficiency and bandwidth of linear polarization converters [11]. They have recently demonstrated metamaterial polarization converters that are capable of rotating the linear polarization to its orthogonal direction over a very broad bandwidth with high efficiency.

For a metamaterial linear polarization converter operating in reflection, their experimental results have shown that over the frequency range from 0.65 to 1.87 terahertz (THz) the cross-polarized reflection carries more than 50% of the incident EM power and the copolarized reflection power is less than 14%. Between 0.73 and 1.8 THz, the cross-polarized reflected power is higher than 80%, and at frequencies near 0.76 THz and 1.36 THz, the copolarized reflected power is less than 1%. Similarly, the same principles applied to polarization conversion in transmission leads to a device with conversion efficiency greater than 50% from 0.52 to 1.82 THz and a maximum efficiency of 80% at 1.04 THz. This design is expected to be applicable to wavelengths ranging from microwaves through infrared light with straightforward scaling of the geometry and appropriate choice of materials.

The metamaterial approaches are versatile and can avoid the restriction of needing materials with intrinsic birefringence or the small optical activity of natural materials. The remaining challenge is then how to further expand the bandwidth and prove the operation of high-efficiency polarimetric devices at a broad range of wavelengths.

Flat optics: Anomalous refraction and reflection

One recent breakthrough in metamaterials is the demonstration of generalized laws of reflection and refraction [12]. The textbook laws state that the direction of travel is determined by both the refractive indices and angle of incidence in the case of refraction and solely on the angle of incidence in the case of reflection. The recent generalized laws expand these concepts to explain how the wave's path changes when there is a phase discontinuity on the interface. When the discontinuity imposes a constant phase gradient from 0 to 2π with uniform energy amplitude on waves propagating through the interface, the outbound wavefront (which is normal to the phase gradient) travels in a direction other than that determined by the traditional laws of reflection and refraction.

Such a constant gradient of phase discontinuity was experimentally realized (identical scattering amplitude of the resonators was also required and realized), and consequently, anomalous reflection and transmission were observed in the mid-infrared [12] and visible [13] regimes. This demonstration may find important applications, such as direction control of light [14], wavefront shaping [15], and flat metalens design [16, 17].

The main challenge with these devices is that the regularly reflected and refracted beams carry most of the incident EM energy, and the intensities of the anomalously reflected and transmitted beams are much weaker than the regular beams. This issue is fundamentally associated with the use of a single-layered metasurface. The anomalous reflection and transmission critically rely on the cross-polarization coupling in anisotropic metamaterials, which is weak for such a single-layered metamaterial.

Using the Fabry-Pérot-like multiple reflection interference and layering principles from their polarization converter design, Antoinette Taylor's team at LANL was able to overcome this limitation on efficiency (see figure 4), demonstrating anomalous refraction with over 50% intensity transmission across most of the 1.0 to 1.4 THz band with transmission peaks over 60%, while the usual refraction direction showed intensity transmissions below 20% and approaching negligible transmission efficiencies at some frequencies [11]. Computational simulations suggest that similar design principles can be used to create a device capable of anomalous reflection.

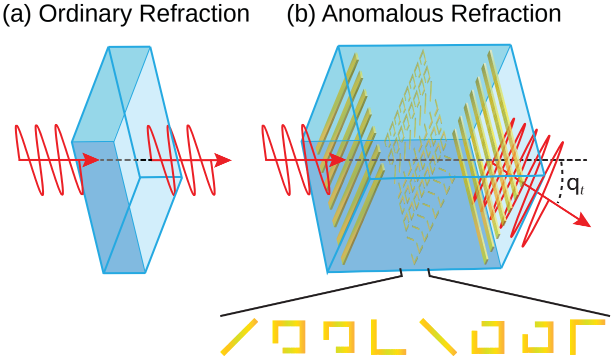


FIGURE 4. (a) This diagram demonstrates ordinary refraction. **(b)** This diagram demonstrates efficient anomalous refraction [11]. The sample consists of a metal wire grating, spacer layer, metamaterial layer designed to impose a phase gradient, a second spacer layer, and a second wire grating, all encapsulated in polyimide.

Going one step beyond simply turning a beam, Vladimir Shalaev's group at Purdue University recently succeeded in creating a flat lens based on exploiting the phase shift from different resonator structures on a planar metamaterial surface [18]. A conventional lens works because its varying thickness creates a phase change across an incoming wave front. Flat lenses instead use an array of resonators with different phase responses to achieve the same feat with a thickness well below the wavelength of the EM wave. By putting resonators with different phase shifts in precisely spaced concentric circles, the researchers were able to reproduce the relative intensity distribution of a lens, although the throughput was only on the order of 10%. Continuing research in this area will lead to more efficient and more specialized designs, providing greater functionality by allowing optical elements to be more easily integrated into devices, especially in applications where mass and volume are an issue.

Metamaterials in wireless antennas

Antennas are crucial elements for microwave wireless communication and wireless devices. The major goals of wireless technology are to reduce the antenna size, increase the radiation efficiency, increase the bandwidth of operation, and increase the gain/directivity. The applications of metamaterials in wireless technology are mainly divided into three categories: (a) where bulk metamaterials are used to enhance the performance of antenna [19–21], (b) where the designs of the antennas are inspired by the unit cell of the metamaterial [22], and (c) where a gain element is integrated into the metamaterial unit cells in the form of self-oscillating metamaterial emitters, or *metamitters* [23].

It is well known that an electrically small electric dipole antenna is an inefficient radiator because it has a very small radiation resistance while simultaneously having a very large capacitive reactance. It thus introduces a large impedance mismatch to any realistic power source and prevents the microwave energy from feeding into the antenna efficiently. To obtain a high overall efficiency, considerable effort must be expended on creating a matching network that forces the total reactance to zero by introducing a very large inductive reactance and tunes the effective input resistance of the antenna to match a 50-ohm source. Generally, this matching method utilizes passive lumped elements. Because of the very large reactance values involved, these matched resonant systems generally have very narrow bandwidths, imperfect efficiencies, and high tolerance requirements for their fabrication.

Metamaterials offer a unique opportunity to match the reactance of the small antenna without a matching network. To achieve impedance matching, the EM properties of the volume adjacent to the antenna are modified using metamaterials concepts. For example, if the antenna is capacitive, then an inductive metamaterial shell is used to make the overall reactance zero. Ziolkowski and Erentok proposed using a spherical shell made out of negative permittivity metamaterials to increase the radiation of the dipole and monopole antenna [19].

Metamaterials also provide an efficient way to obtain directive emission from an antenna. Enoch et al. demonstrated directive emission from a monopole antenna when it was embedded in a metamaterial slab [20], illustrating the ability of metamaterials to change the radiation properties of simple antennas. Using a metamaterial superstrate, rather than fully embedding the antenna, also allows for engineering the gain and bandwidth of an antenna. For example, researchers found that the gain and bandwidth of a microwave patch antenna could be increased by introducing a metamaterial slab on top of a patch antenna [21]. The double-negative-index or negative-index metamaterial coupled to the antenna via a near-field interaction, which increased both the operational bandwidth and antenna gain.

Recently, there have been efforts to make antennas based on a single metamaterial unit cell; such an antenna is termed a *meta-antenna*. In this approach, a single metamaterial element is excited either by a small monopole or by a loop antenna via near-field parasitic coupling. The combined system shows unusually improved radiation efficiency and allows tuning of the input impedance. Ziolkowski et al. experimentally demonstrated a metamaterial-inspired electrically small antenna that reached an overall radiation efficiency of 80% without a matching network [22].

Self-resonating metamaterial antennas, or metamitters, are metamaterial structures in which a source of gain has been integrated into the individual metamaterial elements. For example, a tunnel diode biased into its negative differential resistance regime can be integrated into the gap of an SRR (see figure 5), which acts as both a tank circuit and an efficient antenna [23]. In addition to working as a self-oscillating transmitter, these devices have exhibited a range of nonlinearities, such as frequency mixing, frequency pulling, and bistability, suggesting they may be useful as elements of highly compact detectors. While commercial gallium arsenide (GaAs) tunnel diodes are limited to 12 gigahertz, advanced resonant tunneling diodes will allow similar metamitter designs to operate into the terahertz.

Active metamaterial concepts, discussed below, enable dynamic tuning of the frequency response, beam steering, and variable focusing of the emitted radiation. For example, Kymeta is working toward the commercial production of a metamaterial broadband microwave antenna that uses active metamaterials to electronically steer a radio frequency beam so that it stays locked onto a satellite while in motion without any moving parts [24].

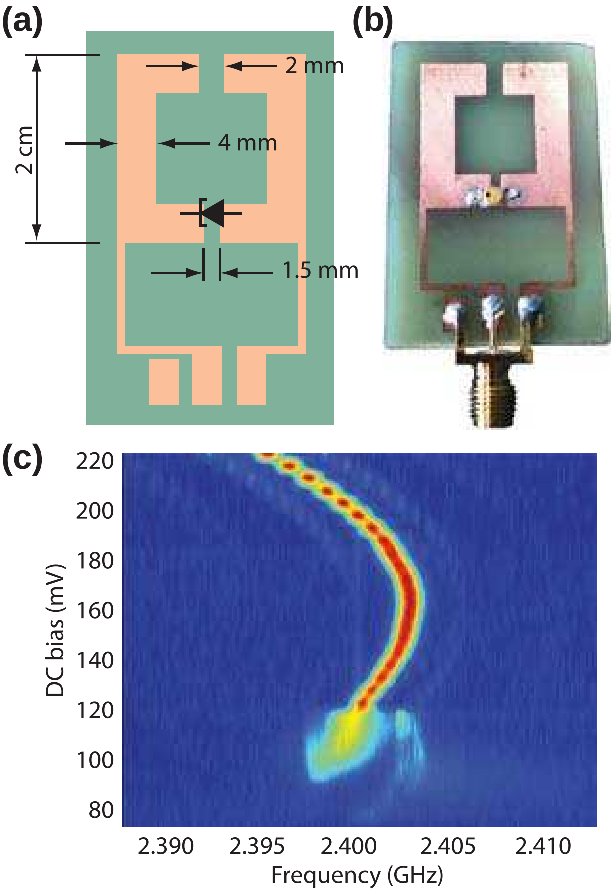


FIGURE 5. (a) This diagram and **(b)** photograph show an individual metamitter element. **(c)** This graph shows the output of the metamitter showing frequency tuning with different applied direct current biases. Image reprinted from [23].

Electrically small antennas will improve the performance of cell phones, personal digital assistants, and Wi-Fi interfaces in laptops. Indeed, a few examples have already shipped in large-volume consumer devices, including Wi-Fi routers made by Netgear. Wireless personal health monitoring systems and miniature wireless sensors for sensor network applications will greatly benefit from the integration of compact efficient antennas and could be a promising commercial application of metamaterials.

Active control of metamaterials

In general, after a metamaterial structure is fabricated on a substrate, its resonance strength, frequency, and the relative phase of the individual elements are already fixed. Due to the highly dispersive properties, the operational bandwidth of metamaterials is often very narrow; this static nature makes it beneficial for some applications but difficult to use for broadband or tunable applications. Thus, dynamically and actively tunable metamaterials are highly desirable to enhance functionality.

Metamaterials derive their behavior from combinations of effective permittivity and effective permeability that arise as a consequence of averaging over the behavior of a set of meta-atoms. It is frequently useful to view this subwavelength meta-atom as an inductive-capacitive circuit whose resonance frequency and amplitude are determined by the effective capacitance and the inductance provided by the unit cell. Therefore, tunability can be obtained by changing the values of the capacitance and/or inductance through extrinsic or intrinsic stimuli. However, directly changing the properties of metallic elements is difficult except in a few exotic cases, such as when a superconductor or graphene is being used instead of a normal metal.

Instead, dynamic metamaterials are usually obtained by integrating semiconductors or a dielectric material whose properties can be altered either through optical or electrical excitation [25]. For example, changing the metamaterial substrate affects both the resonance strength (via substrate losses) and the resonance frequency (via the substrate's dielectric constant). The functionality in active/dynamic metamaterials is essentially determined by modifying the metamaterial substrate or by incorporating materials into critical regions of the resonant elements.

Active metamaterials are of particular interest at terahertz frequency from the device point of view. Researchers at LANL demonstrated many of these concepts for realistic device applications by integrating semiconductors in the metamaterial designs [25]. Optical illumination was used to dynamically change the resonance amplitude by photoexciting the carriers in the metamaterial's unit cell, which damped the resonance because of the increased loss in the capacitor gap. In another demonstration, Hou-Tong Chen demonstrated the frequency redshifting in a metamaterial that had silicon semiconductor pads; upon photoexcitation, the pads became conductive and thus increased the overall capacitance. Very recently, a similar approach was demonstrated by another group to show the resonance blueshifting by using optical excitation [26].

Another approach to control the EM properties of individual meta-atoms remotely using light is to integrate varactor diodes into each meta-atom, which could then be controlled by a nearby light-emitting diode (LED) [27]. In the absence of light, incoming microwaves would reflect off this SRR array like a flat mirror. By increasing the brightness of selected LEDs, the angle of reflection could be altered. The array could even focus or defocus microwaves, as if it were a parabolic mirror.

Electrically controllable terahertz metamaterials, shown in figure 6, were demonstrated by Hou-Tong Chen at LANL [28]. Room-temperature electrically switchable metamaterials were first created by fabricating planar metallic metamaterials on a thin n-doped GaAs layer. The gold SRRs and n-GaAs form a Schottky diode structure, enabling control of the charge carriers in the metamaterial split gaps by application of a reverse voltage bias, which in turn tunes the resonant amplitudes. This technique has been applied to realize terahertz electrical modulators and phase shifters. Tunable terahertz metamaterials have also been realized using a microelectromechanical systems structure.

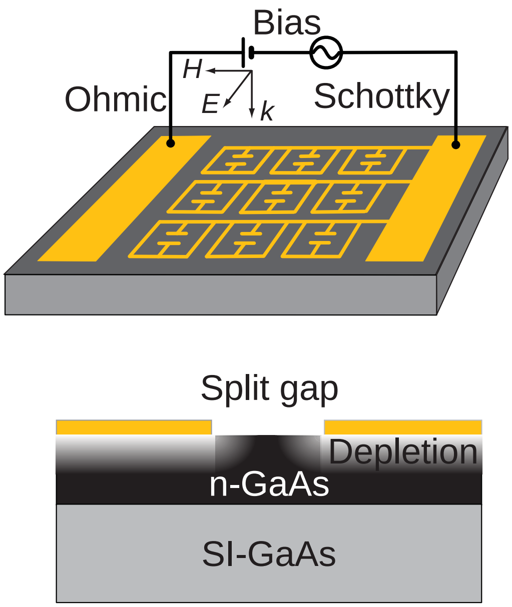


FIGURE 6. Electrically switchable terahertz metamaterials were first created by fabricating planar metallic metamaterials on a thin n-doped GaAs layer. The gold SRRs and n-GaAs form a Schottky diode structure, enabling control of the charge carriers in the metamaterial split gaps by application of a reverse voltage bias, which in turn tunes the resonant amplitudes. Image is adapted from [28].

Currently, the terahertz technology is suffering from the lack of compact sources and detectors. However, the development of the quantum cascade terahertz laser might be the first-generation terahertz device that might benefit from the metamaterial-based active terahertz modulators. In this case, the modulator could be designed to match the frequency of terahertz radiation. At optical frequencies, reconfiguration of negative-index metamaterials has been proposed and designed by controlling the magnetic resonance via tuning the permittivity of the embedded anisotropic liquid crystals [29]. Such a structure has recently been experimentally investigated by infiltrating fishnet metamaterials with nematic liquid crystals [30]. Experimental results showed a significant change in the optical transmission with a moderate laser power.

A group of researchers led by Antoinette Taylor at LANL demonstrated ultrafast switching of the negative-index optical metamaterials using all optical switching [31]. Their device consisted of metallic fishnet structures with a thin amorphous silicon layer in between. The optical excitation allowed photo-induced carrier injection, which dynamically tuned the resonance behavior of the metamaterials. The device was able to modulate at the communication wavelength with a speed of one terabit per second. The planar design of such a device can be fabricated easily using the existing deep ultraviolet photolithography technique.

A significant portion of the metamaterial research is committed to the development of such active devices. While advantageous for some applications, the significant energy loss and narrow bandwidth of operation of metamaterials are a detriment for many other applications. The lack of compact devices has always been a roadblock for practical applications of terahertz technology, indicating that this area will likely benefit most from the emergence of active metamaterials devices. For example, these kinds of terahertz modulators might emerge as an integrated part of resonant tunneling diode-based terahertz sources. The fabrication of active terahertz modulators is compatible with current high-volume microfabrication methods; therefore, commercialization does not require any extra infrastructure development. An electrically controllable phase modulator is another metamaterial application that is likely to enable the operation of terahertz scanners without moving parts in the midterm (i.e., 4–8 year) timescale.

Conclusions

Metamaterials are a promising technology on the verge of transitioning from pure laboratory research to commercial applications. Metamaterial absorbers are likely to have a significant effect on the areas of EM signature manipulation, solar energy, and thermophotovoltaics. Electrically small antennas, artificial ferrites, and metamitters will have a significant effect on the miniaturization of wireless communications devices. Metamaterial-based, ultrathin, lightweight optics will affect areas ranging from radar to terahertz and infrared imaging and possibly communications, optical microscopy, and lithography.

In addition to the discussed applications, electromagnetic metamaterials have inspired analogous efforts to control other wave phenomena, including the emerging field of acoustic metamaterials. Preliminary demonstrations indicate that acoustic metamaterials may lead to significant advances in ultrasound and sonar imaging resolution, sound isolation, and acoustical cloaking. In summary, metamaterials are a rapidly advancing, dynamic area of research with often surprising discoveries routinely emerging. In some areas, notably compact antennas, metamaterials are rapidly maturing into a commercially relevant technology.

About the authors

Jane E. Heyes is a research technologist on Antoinette Taylor's ultrafast optics team at the Center for Integrated Nanotechnologies at Los Alamos National Laboratory (LANL). She holds a bachelor's and a master's degree in electrical engineering from Stanford University.

Nathaniel Grady received his PhD in applied physics from Rice University in 2010. His doctoral work spanned a range of experimental and theoretical investigations into the fundamental nature and application of plasmons in metallic nanostructures including linear and nonlinear spectroscopy of individual nanoparticles, surface-enhanced spectroscopy, surface-enhanced fluorescence, nanoparticle-enhanced photovoltaics, and solar-thermal energy harvesting. Following his PhD, he studied the guiding of light on the nanoscale using plasmons propagating on metal nanowires as a postdoctoral student at the Institute of Physics, Chinese Academy of Sciences in Beijing. He is currently a postdoctoral student at LANL where he is studying terahertz, microwave, and infrared metamaterials for ultrathin flat optics, the generation of high-intensity terahertz waves, and nonlinear terahertz spectroscopy.

Dr. Diego Dalvit is a technical staff scientist in the Theory Division of LANL. He leads the LANL theory team working on modeling and simulation of light-matter interactions in metamaterials, nanophotonics, and Casimir physics. His other areas of expertise are in quantum information science and technology, including decoherence, measurement, and control of open quantum systems. He has been a visiting scholar at the Ecole Normale Superieure and the French National Center for Scientific Research at the German Academic Exchange Service. Dalvit has also been a LANL director's postdoctoral fellow.

Dr. Antoinette (Toni) Taylor is the leader of the Materials Physics and Applications Division at LANL. Prior to this position, she was director of the Center for Integrated Nanotechnologies, a joint Sandia/LANL Nanoscience Research Center funded through the Office of Basic Energy Sciences. Her research interests include the investigation of ultrafast dynamical nanoscale processes in materials, the development of novel optical functionality using metamaterials, and the development of novel optics-based measurement techniques for the understanding of new phenomena. She has published over 300 papers in these areas, written three book chapters, and edited five books. She is a former director-at-large of the Optical Society of America (OSA), topical editor of the *Journal of the Optical Society B: Optical Physics,* and a member of the National Academies' Board of Physics and Astronomy Solid State Science Committee. Taylor has also chaired the National Academies' Committee on Nanophotonics Applicability and Accessibility. Currently, she is a Chair of the Division of Laser Science of the American Physical Society, and the OSA representative on the Joint Council of Quantum Electronics. She is a LANL laboratory fellow and a fellow of the American Physical Society, the OSA, and American Association for the Advancement of Science. In 2003, Taylor won the inaugural Los Alamos Fellow's Prize for Outstanding Leadership in Science and Engineering.

References

[1] Veselago VG. "The electrodynamics of substances with simultaneously negative values of ε and μ." *Soviet Physics Uspekhi.* 1968;10(4):509. doi: 10.1070/PU1968v010n04ABEH003699.

[2] Pendry JB, Holden AJ, Robbins DJ, Stewart WJ. "Magnetism from conductors and enhanced nonlinear phenomena." *IEEE Transactions on Microwave Theory and Techniques.* 1999;47(11):2075–2084. doi: 10.1109/22.798002.

[3] Shelby RA, Smith DR, Schultz S. "Experimental verification of a negative index of refraction." *Science.* 2001;292(5514):77–79. doi: 10.1126/science.1058847.

[4] Soukoulis CM, Wegener M. "Past achievements and future challenges in the development of three-dimensional photonic metamaterials." *Nature Photonics.* 2011;5(9):523–530. doi: 10.1038/nphoton.2011.154.

[5] Landy NI, Sajuyigbe S, Mock JJ, Smith DR, Padilla WJ. "Perfect metamaterial absorber." *Physical Review Letters.* 2008;100(20):207402. doi: 10.1103/PhysRevLett.100.207402.

[6] Chen H. "Interference theory of metamaterial perfect absorbers." *Optics Express.* 2012;20(7):7165–7172. doi: 10.1364/OE.20.007165.

[7] Aydin K, Ferry VE, Briggs RM, Atwater HA. "Broadband polarization-independent resonant light absorption using ultrathin plasmonic super absorbers." *Nature Communications.* 2011;2(10):517. doi: 10.1038/ncomms1528.

[8] Liu X, Tyler T, Starr T, Starr AF, Jokerst NM, Padilla WJ. "Taming the blackbody with infrared metamaterials as selective thermal emitters." *Physical Review Letters.* 2011;107(4):045901. doi: 10.1103/PhysRevLett.107.045901–045905.

[9] Kirchhoff G. "Über das Verhältnis zwischen dem Emissionsvermöogen und dem Absorptionsvermögen der Körper für Wärme und Licht." *Annalen der Physik und Chemie.* 1860;109:275–301. doi: 10.1002/andp.18601850205. [English translation, "On the relation between the radiating and the absorbing powers of different bodies for light and heat." *Philosophical Magazine.* 1860;4(20):1–21].

[10] Kim Y, Lin S, Chang ASP, Lee J, Ho K. "Analysis of photon recycling using metallic photonic crystal." *Journal of Applied Physics.* 2007;102(6):063107. doi: 10.1063/1.2779271.

[11] Grady N, Heyes J, Chowdhury DR, Zeng Y, Taylor AJ, Dalvit DAR, Chen H. "Terahertz metamaterials for linear polarization conversion and anomalous refraction." *Science.* 2013;340(6138):1304–1307. doi: 10.1126/science.1235399.

[12] Yu N, Genevet P, Kats MA, Aieta F, Tetienne J, Capasso F, Gaburro Z. "Light propagation with phase discontinuities: Generalized laws of reflection and refraction." *Science.* 2011;334(6054):333–337. doi: 10.1126/science.1210713.

[13] Ni X, Emani NK, Kildishev AV, Boltasseva A, Shalaev VM. "Broadband light bending with plasmonic nanoantennas." *Science.* 2012;335(6067):427. doi: 10.1126/science.1214686.

[14] Aieta F, Genevet P, Yu N, Kats MA, Gaburro Z, Capasso F. "Out-of-plane reflection and refraction of light by anisotropic optical antenna metasurfaces with phase discontinuities." *Nano Letters.* 2012;12(3):1702–1706. doi: 10.1021/nl300204s.

[15] Genevet P, Yu N, Aieta F, Lin J, Kats MA, Blanchard R, Scully MO, Gaburro Z, Capasso F. "Ultra-thin plasmonic optical vortex plate based on phase discontinuities." *Applied Physics Letters.* 2012;100(1):013101. doi: 10.1063/1.3673334.

[16] Chen X, Huang L, Mühlenbernd H, Li G, Bai B, Tan Q, Jin G, Qiu C, Zhang S, Zentgraf T. "Dual-polarity plasmonic metalens for visible light." *Nature Communications.* 2012;3(11):1198. doi: 10.1038/ncomms2207.

[17] Aieta F, Genevet P, Kats MA, Yu N, Blanchard R, Gaburro Z, Capasso F. "Aberration-free ultrathin flat lenses and axicons at telecom wavelengths based on plasmonic metasurfaces." *Nano Letters.* 2012;12(9):4932–4936. doi: 10.1021/nl302516v.

[18] Ni X, Ishii S, Kildishev AV, Shalaev VM. "Ultra-thin, planar, Babinet-inverted plasmonic metalenses." *Light: Science & Applications.* 2013;2(4):e72. doi: 10.1038/lsa.2013.28.

[19] Ziolkowski RW, Erentok A. "Metamaterial-based efficient electrically small antennas." *IEEE Transactions on Antennas and Propagation.* 2006;54(7):2113–2130. doi: 10.1109/TAP.2006.877179.

[20] Enoch S, Tayeb G, Sabouroux P, Guérin N, Vincent P. "A metamaterial for directive emission." *Physical Review Letters.* 2002;89(21):213902. doi: 10.1103/PhysRevLett.89.213902.

[21] Ju J, Kim D, Lee WJ, Choi JI. "Wideband high-gain antenna using metamaterial superstrate with the zero refractive index." *Microwave and Optical Technology Letters.* 2009;51(8):1973–1976. doi: 10.1002/mop.24469.

[22] Ziolkowski RW, Peng J, Nielsen JA, Tanielian MH, Holloway CL. "Experimental verification of Z antennas at UHF frequencies." *IEEE Antennas and Wireless Propagation Letters.* 2009;8:1329–1333. doi: 10.1109/LAWP.2009.2038180.

[23] O'Hara JF, Reiten MT, Colestock P, Earley L, Taylor A. "Tunnel-diode loaded split-ring resonators as a foundation for nonlinear metamaterials." *Proceedings of the SPIE.* 2011;8093:809304. doi: 10.1117/12.891402.

[24] For more information, visit: http://www.kymetacorp.com/.

[25] Chen H, O'Hara JF, Azad AK, Taylor AJ. "Manipulation of terahertz radiation using metamaterials." *Laser & Photonics Reviews.* 2011;5(4):513–533. doi: 10.1002/lpor.201000043.

[26] Shen N, Massaouti M, Gokkavas M, Manceau J, Ozbay E, Kafesaki M, Koschny T, Tzortzakis S, Soukoulis CM. "Optically implemented broadband blueshift switch in the terahertz regime." *Physical Review Letters.* 2011;106(3):037403. doi: 10.1103/PhysRevLett.106.037403.

[27] Shadrivov IV, Kapitanova PV, Maslovski SI, Kivshar YS. "Metamaterials controlled with light." *Physical Review Letters.* 2012;109(8):083902. doi: 10.1103/PhysRevLett.109.083902.

[28] Chen H, Padilla WJ, Zide JMO, Gossard AC, Taylor AJ, and Averitt RD. "Active terahertz metamaterial devices." *Nature.* 2006;444(7119):597–600. doi: 10.1038/nature05343.

[29] Wang X, Kwon D, Werner DH, Khoo I, Kildishev AV, Shalaev VM. "Tunable optical negative-index metamaterials employing anisotropic liquid crystals." *Applied Physics Letters.* 2007;91(14):143122. doi: 10.1063/1.2795345.

[30] Minovich A, Farnell J, Neshev DN, McKerracher I, Karouta F, Tian J, Powell DA, Shadrivov IV, Tan HH, Jagadish C, Kivshar YS. "Liquid crystal based nonlinear fishnet metamaterials." *Applied Physics Letters.* 2012;100(12):121113. Available at: http://dx.doi.org/10.1063/1.3695165.

[31] Dani KM, Ku Z, Upadhya PC, Prasankumar RP, Brueck SRJ, Taylor AJ. "Subpicosecond optical switching with a negative index metamaterial." *Nano Letters.* 2009;9(10):3565–3569. doi: 10.1021/nl9017644.

The Next Wave is published to disseminate technical advancements and research activities in telecommunications and information technologies. Mentions of company names or commercial products do not imply endorsement by the US Government.

This publication is available online at http://www.nsa.gov/research/tnw/index.shtml. For more information, please contact us at TNW@tycho.ncsc.mil.

NATIONAL SECURITY AGENCY UNITED STATES OF AMERICA

CENTRAL SECURITY SERVICE UNITED STATES OF AMERICA

RESEARCH DIRECTORATE ADVANCING INTELLIGENCE THROUGH SCIENCE

Securing the cloud with homomorphic encryption

Research Directorate staff

The word *homomorphic* has roots in Greek and loosely translates as "same shape" or "same form." In relation to cryptography, the concept is that operations can be performed on encrypted data without sharing the secret key needed to decrypt the data. Homomorphic encryption has great utility in cloud computing, particularly for those that wish to house encrypted data on cloud providers' servers.

A major hurdle to the adoption of cloud-based services is security. Cloud users, particularly at the enterprise and government level, are concerned with losing control of, or just plain losing, their data once it is placed in the cloud. The abstractness of cloud storage makes it difficult for consumers to feel comfortable that their data is well protected by cloud service providers. Encryption could alleviate this issue. However, if you want to manipulate your encrypted data in the cloud, the secret key to decrypt your data must be shared with the cloud provider. This sort of defeats the idea of a secret key. Sharing this key of course would allow the current cloud provider (or future provider if the service changes hands) access to your data. The answer to this problem could be homomorphic encryption.

For example, a bakery in New York that uses a cloud service provider's infrastructure to host their e-mail wants to search through those e-mails for an order erroneously sent to Hoboken, New Jersey. If the data is plaintext, the subscriber just plugs in a search term (e.g., "cupcakes Hoboken") and views the results. If the data is encrypted, the bakery will need to share the secret key with the cloud provider to access the information stored on the provider's servers to query against the data. Sharing that secret key now potentially gives the provider access to the company's data, and if there is a security breach, it may also give cybercriminals access to the data. Homomorphic encryption would allow the bakery's owners to search the encrypted e-mails for items related to the Hoboken mishap and get results as if querying against the plaintext data, without sharing the key.

The idea of homomorphic encryption has been around for about 30 years, and thanks to a significant breakthrough in 2009, the end game of a practical fully homomorphic encryption solution is in sight. There are fully homomorphic encryption solutions that exist today, but because of limitations related mainly to the complexity of computations, these solutions are not considered practical for use with today's applications. These limitations are being addressed, and some say a practical solution could be achieved within a decade. If a practical, fully homomorphic solution can be created, it could be the catalyst that breaks down the security barrier to widespread cloud adoption.

Technical overview

Fully versus somewhat homomorphic encryption

There are two types of homomorphic encryption: fully homomorphic encryption (FHE) and somewhat homomorphic encryption (SHE). Each type differs in the number of operations that can be performed on encrypted data. FHE allows for an unlimited, arbitrary number of computations (both addition and multiplication) to be performed on encrypted data. SHE cryptosystems support a limited number of operations (i.e., any amount of addition, but only one multiplication) and are faster and more compact than FHE cryptosystems [1].

Bootstrapping and lattices

One of the hindrances limiting the feasibility of FHE is managing the so called "noise." Noise, in this case, refers to the distortion of ciphertexts (i.e., encoded text) that occurs after each operation (e.g., addition or multiplication) is performed. As more and more additions and multiplications are performed, the noise level becomes too high, and the resulting ciphertexts become indecipherable. Ciphertexts can be refreshed easily by decrypting them, but the idea behind homomorphic encryption is to not share the secret key required to do the decryption.

Craig Gentry used a process called bootstrapping to overcome this noise problem in SHE solutions. Bootstrapping modifies an SHE solution so it can homomorphically run its own decryption procedure by adding an encryption of the secret key to the public key (see figure 1). This is accomplished by using a sparse subset-sum problem (SSSP) or augmenting the public key with a large set of vectors so that a sparse subset of the vectors will add up to be the secret key.

The idea of bootstrapping calls for double encrypting the data and, as processes run, removing a layer of encryption. Gentry's bootstrapping procedure adds another layer of encryption after a few computations using an encrypted key to unlock the inner layer of scrambling. This process "refreshes" the still-encrypted data and could allow for an infinite number of computations, effectively turning an SHE solution into an FHE solution. However, each extra layer of encryption will increase the overall computational effort needed to complete a query [1, 2, 3].

The downside of Gentry's bootstrapping idea is that it requires huge amounts of computational effort. For example, if the process were to be used by Google to search the web homomorphically, the normal computing time would be multiplied

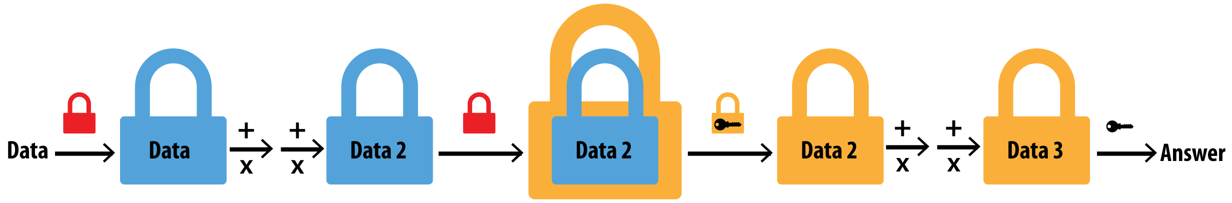


FIGURE 1. Gentry's bootstrapping method modifies an SHE solution so it can homomorphically run its own decryption procedure by adding an encryption of the secret key to the public key.

by about a trillion, according to Gentry. This extra computing time is one of the reasons that a practical FHE solution is not available for implementation today. Although strides are being made every day to overcome the amount of processing overhead needed to use these solutions, the schemes tend to be difficult to understand and even harder to implement. Also, the rate of improvement that is occurring in this field could make it hard for early adopters to keep up with the pace of innovation [2, 4].

Another important aspect of FHE solutions is their basis on ideal lattices. Ideal lattices are special classes of lattices that are particularly useful in cryptography [5, 6]. Lattice-based encryption schemes are the focus of FHE solutions because they have simple decryption algorithms which could lessen the computational overhead associated with bootstrapping SHEs [7]. Lattice-based schemes are also attractive for FHEs because they are based on worst-case hardness—meaning that there is a very small chance of attacks succeeding against lattice-based schemes [6]. In this regard, lattice-based cryptographic schemes are believed to even be secure against attacks using quantum computers [6].

Trends

As cloud services spread globally, the need for an FHE scheme will become more important and may begin to draw interest from entities using or planning to use the cloud (public or private) to store data. A report from the International Data Corporation (IDC) on global public cloud-enabling IT infrastructure forecasts that from 2013 through 2017, the majority of cloud adopters will be located in the US, followed closely by subscribers in Western Europe. The US dominance in this area is attributed to the availability of reasonably priced broadband access and the fact that most first-to-market cloud services were located in the US. However, IDC sees adoption in emerging markets (especially Asia) exhibiting strong growth with a more widespread distribution of adoption across all regions occurring beyond 2014 [8].

Two US government entities, the Defense Advanced Research Projects Agency (DARPA) and the Intelligence Advanced Research Projects Activity (IARPA), issued Broad Agency Announcements (BAAs) for a solution that could perform computations on encrypted data (i.e., homomorphic encryption). In April 2011, DARPA awarded approximately $5 million to Galois, Inc. to be the research integrator for the Programming Computation on Encrypted Data (PROCEED) program. This award was part of a five-year effort by DARPA worth a total of $20 million.

In December 2010, IARPA issued a request for proposals for a program called Security And Privacy Assurance Research (SPAR). The goal of the both programs is to make it feasible to execute programs on encrypted data without having to decrypt the data first. DARPA's stated goal was to reduce the computing time for an FHE solution by a factor of 10 million [9].

Conclusion

A practical FHE solution would see widespread use by cloud service providers, significantly hardening cloud security and making cloud storage a more viable option for consumers. Some have predicted that, thanks to Gentry's revelation and the momentum that it generated in the world of cryptography, an FHE solution may be feasible in another decade [2]. A Chief Technology Officer in the UK noted that, by applying Moore's law, it would take 40 years before a fully homomorphic search would be as efficient as a Google search today [10].

Researchers worldwide are actively engaged in trying to perfect a practical FHE solution. Recent breakthroughs include a homomorphic encryption scheme from Fujitsu using batch encryption vice the bit-level encryption usually seen in FHE solutions [11]. Also, in June 2013, researchers from the Massachusetts Institute of Technology, the University of Toronto, and Microsoft Research created a three-part encryption scheme that uses homomorphic encryption as well as two other cryptographic techniques (i.e., garbled circuits and attribute-based encryption) [12]. Although these solutions are not commercially available (Fujitsu hopes to market its solution by 2015), they will perpetuate the continued efforts to field a fully functioning FHE solution.

References

[1] Hayes B. “Alice and Bob in cipherspace.” *American Scientist.* 2012;100(5). Available at: http://www.americanscientist.org/issues/pub/2012/5/alice-and-bob-in-cipherspace/5.

[2] Greenberg A. “IBM’s blindfolded calculator.” *Forbes.* Jul 2009. Available at: http://www.forbes.com/forbes/2009/0713/breakthroughs-privacy-super-secret-encryption.html.

[3] Gentry C, Halevi S. “Implementing Gentry’s fully-homomorphic encryption scheme.” *Cryptology ePrint Archive.* 2011. Available at: http://eprint.iacr.org/2010/520.

[4] Aguilar-Melchor C, Fau S, Fontaine C, Gogniat G, Sirdey R. “Recent advances in homomorphic encryption.” *IEEE Signal Processing Magazine.* 2013. doi: 10.1109/MSP.2012.2230219.

[5] “Ideal lattice cryptography.” *Wikipedia* [last modified 2013 Jan 10]. Available at: http://en.wikipedia.org/wiki/Ideal_lattice_cryptography.

[6] “Lattice-based cryptography.” *Wikipedia* [last modified 2013 Jan 16]. Available at: http://en.wikipedia.org/wiki/Lattice-based_cryptography.

[7] Gentry C. “A fully homomorphic encryption scheme” [dissertation]. Stanford University; 2009. Available at: http://crypto.stanford.edu/craig/craig-thesis.pdf.

[8] Turner MJ, Villars RL, Scaramella J, Mehra R, DuBois L, Iacono D, Gillen A, Chen G, Grady J, Grieser T, Eastwood M. “Worldwide public cloud enabling IT infrastructure 2013–2017 forecast.” International Data Corporation. 2013. Doc # 240635.

[9] Greenberg A. “DARPA will spend $20 million to search for crypto’s holy grail.” *Forbes.* Apr 2011. Available at: http://www.forbes.com/sites/andygreenberg/2011/04/06/darpa-will-spend-20-million-to-search-for-cryptos-holy-grail/.

[10] Schneier B. “Homomorphic encryption breakthrough.” *Schneier on Security* [blog]. Jul 2009. Available at: https://www.schneier.com/blog/archives/2009/07/homomorphic_enc.html.

[11] Fujitsu Laboratories Ltd. “Fujitsu develops world’s first homomorphic encryption technology that enables statistical calculations and biometric authentication” [press release]. Aug 2013. Available at: http://www.fujitsu.com/global/news/pr/archives/month/2013/20130828-01.html.

[12] Hardesty L.“Securing the cloud.” *MIT news.* Jun 2013. Available at: http://web.mit.edu/newsoffice/2013/algorithm-solves-homomorphic-encryption-problem-0610.html .

The Next Wave is published to disseminate technical advancements and research activities in telecommunications and information technologies. Mentions of company names or commercial products do not imply endorsement by the US Government.

This publication is available online at http://www.nsa.gov/research/tnw/index.shtml. For more information, please contact us at TNW@tycho.ncsc.mil.

AT A GLANCE

Technology forecasting

ELECTROMAGNETIC METAMATERIALS

The limitations of natural materials are a major obstacle that must be overcome to meet the ever-increasing demand for faster, lighter, cheaper, and more compact devices. By providing effective material properties not found in nature, electromagnetic metamaterials have the potential to aid in the creation of ultrathin planar lenses, superresolution microscopes, compact antennas, faster computer chips, and surfaces that radically alter or cloak the electromagnetic signature of an object.

TRANSISTORS

Transistors are the building blocks to computer chips; as they get smaller, computers get faster and more energy efficient. But transistors and the copper wires that connect to them can only shrink so much before they reach the limits of quantum mechanics. Researchers are making strides in other fields to solve this problem and find new methods.

HOMOMORPHIC ENCRYPTION

Homomorphic encryption allows operations to be performed on encrypted data without sharing the secret key needed to decrypt the data. If a practical, fully homomorphic solution can be created, it could be the catalyst that breaks down the security barrier to widespread cloud adoption.

SUPERCONDUCTIVE ELECTRONICS

The energy dissipation of complementary metal-oxide semiconductor transistors is reaching physical limits and has become a difficult barrier to building more powerful supercomputers. Digital integrated circuits based on superconductive single-flux-quantum (SFQ) logic offers a combination of high-speed and ultralow power dissipation unmatched by any other device.

PLASMONIC INTERCONNECTS

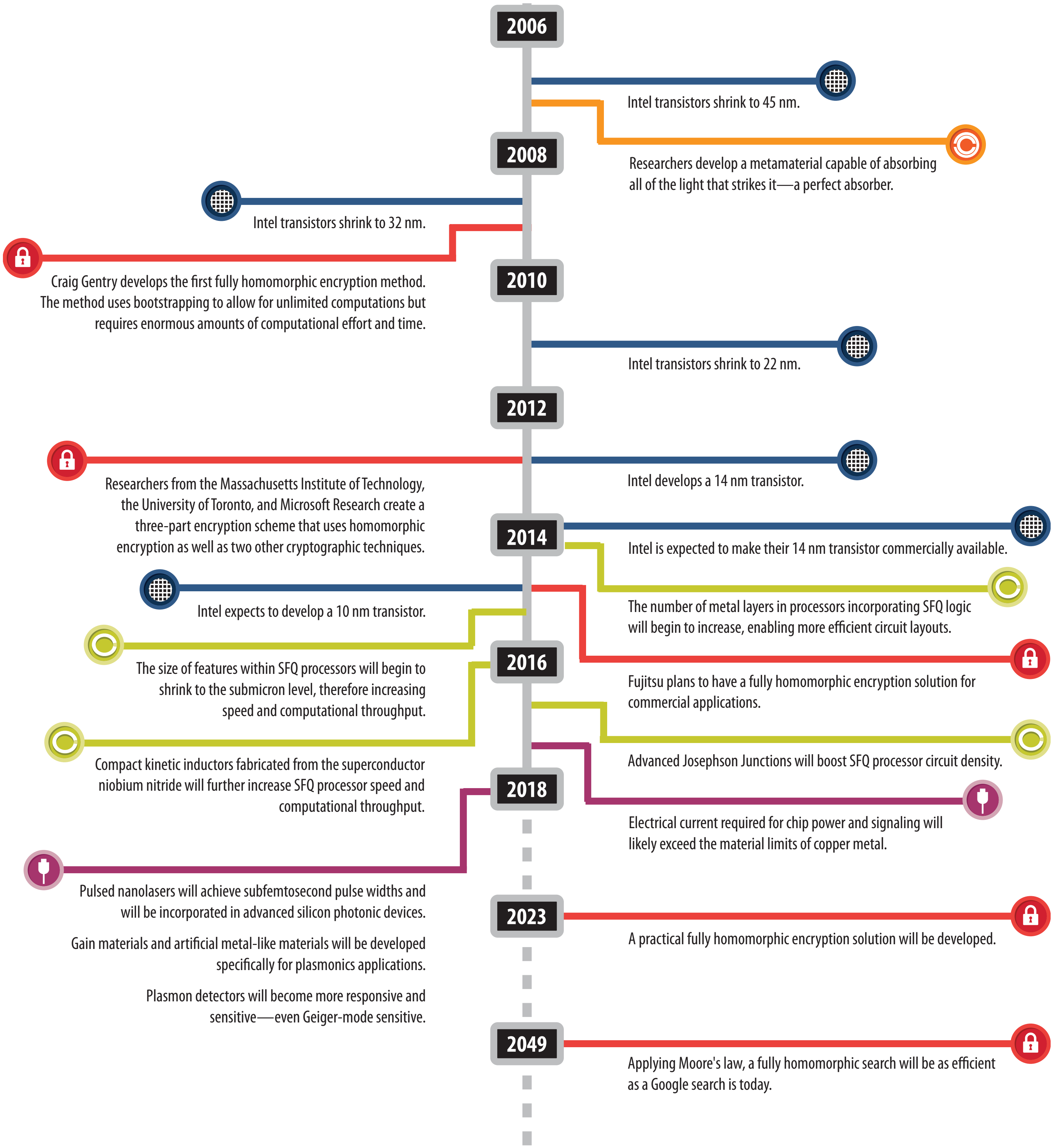
Plasmonic interconnects are composed of optical fields bound to metal surfaces; they offer the benefits of high-bandwidth signaling with physical confinement down to the true nanoscale due to the metals on which plasmons are supported. Plasmonic interconnects have the potential to alleviate the current communications bottleneck in computer architectures.

2000

David Smith and his colleagues demonstrate composite metamaterials, using a combination of plasmonic-type metal wires and a split-ring resonator array to create a negative permittivity and negative permeability in the microwave regime. They demonstrate a negative index of refraction—this begins the study of metamaterials.

2002

2004



2006

Intel transistors shrink to 45 nm.

Researchers develop a metamaterial capable of absorbing all of the light that strikes it—a perfect absorber.

2008

Intel transistors shrink to 32 nm.

Craig Gentry develops the first fully homomorphic encryption method. The method uses bootstrapping to allow for unlimited computations but requires enormous amounts of computational effort and time.

2010

Intel transistors shrink to 22 nm.

2012

Researchers from the Massachusetts Institute of Technology, the University of Toronto, and Microsoft Research create a three-part encryption scheme that uses homomorphic encryption as well as two other cryptographic techniques.

Intel develops a 14 nm transistor.

2014

Intel is expected to make their 14 nm transistor commercially available.

The number of metal layers in processors incorporating SFQ logic will begin to increase, enabling more efficient circuit layouts.

Intel expects to develop a 10 nm transistor.

The size of features within SFQ processors will begin to shrink to the submicron level, therefore increasing speed and computational throughput.

Fujitsu plans to have a fully homomorphic encryption solution for commercial applications.

2016

Compact kinetic inductors fabricated from the superconductor niobium nitride will further increase SFQ processor speed and computational throughput.

Advanced Josephson Junctions will boost SFQ processor circuit density.

Electrical current required for chip power and signaling will likely exceed the material limits of copper metal.

2018

Pulsed nanolasers will achieve subfemtosecond pulse widths and will be incorporated in advanced silicon photonic devices.

Gain materials and artificial metal-like materials will be developed specifically for plasmonics applications.

Plasmon detectors will become more responsive and sensitive—even Geiger-mode sensitive.

2023

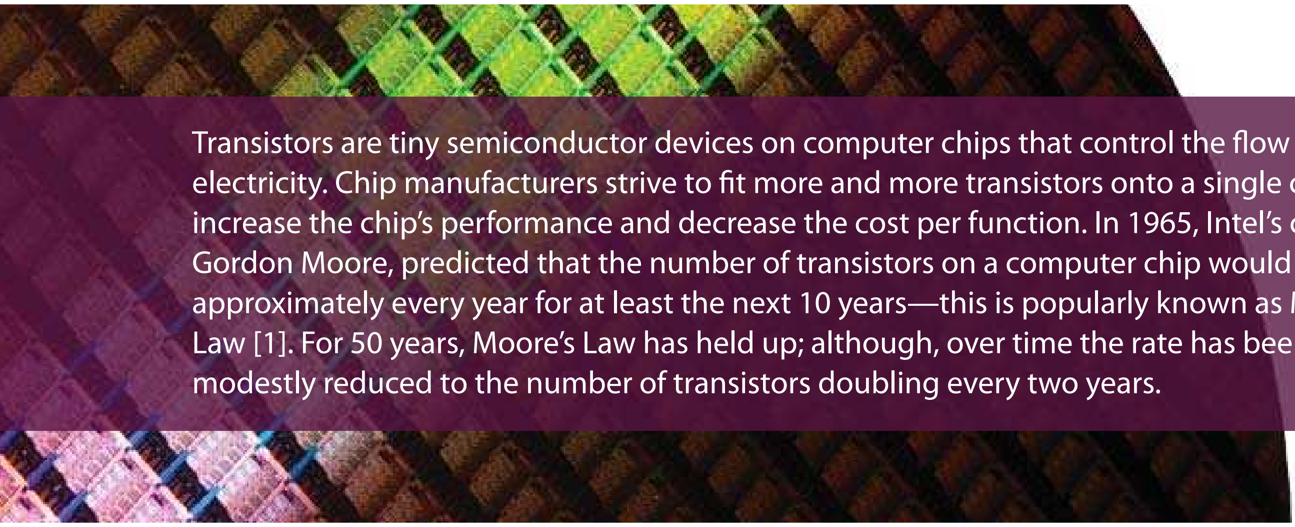
A practical fully homomorphic encryption solution will be developed.

2049

Applying Moore's law, a fully homomorphic search will be as efficient as a Google search is today.

POINTERS

Today’s Tiny Transistors



Transistors are tiny semiconductor devices on computer chips that control the flow of electricity. Chip manufacturers strive to fit more and more transistors onto a single chip to increase the chip’s performance and decrease the cost per function. In 1965, Intel’s cofounder, Gordon Moore, predicted that the number of transistors on a computer chip would double approximately every year for at least the next 10 years—this is popularly known as Moore’s Law [1]. For 50 years, Moore’s Law has held up; although, over time the rate has been modestly reduced to the number of transistors doubling every two years.

In 1971, Intel’s first computer chip contained 2,300 transistors, each measuring 10,000 nanometers (nm)—slimmer than a strand of human hair [2]. Today, Intel’s featured processor chip has 1.4 billion tri-gate transistors, each measuring 22 nm [3] (see figure 1). Intel’s road map shows that they have developed a 14 nm transistor and will begin working on a 10 nm transistor in 2015 [4].

Intel has kept up with Moore's Law (see figure 2); however, as transistors get smaller, they present more challenges, such as current leakage. Eventually, transistor size will reach its physical limits—what then? How will we be able to increase the computation and memory capacity? The research discussed in this issue answers just that.

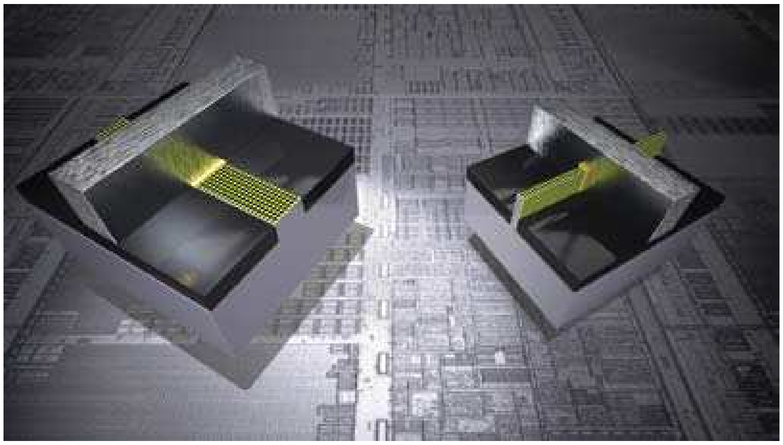


FIGURE 1. This illustration compares Intel’s 32 nm transistor to their 22 nm transistor. On the left side is the 32 nm planar transistor in which the current (represented by the yellow dots) flows in a plane underneath the gate. On the right is the 22 nm three-dimensional tri-gate transistor with current flowing on three sides of a vertical fin. [Image provided by Intel Corporation.]

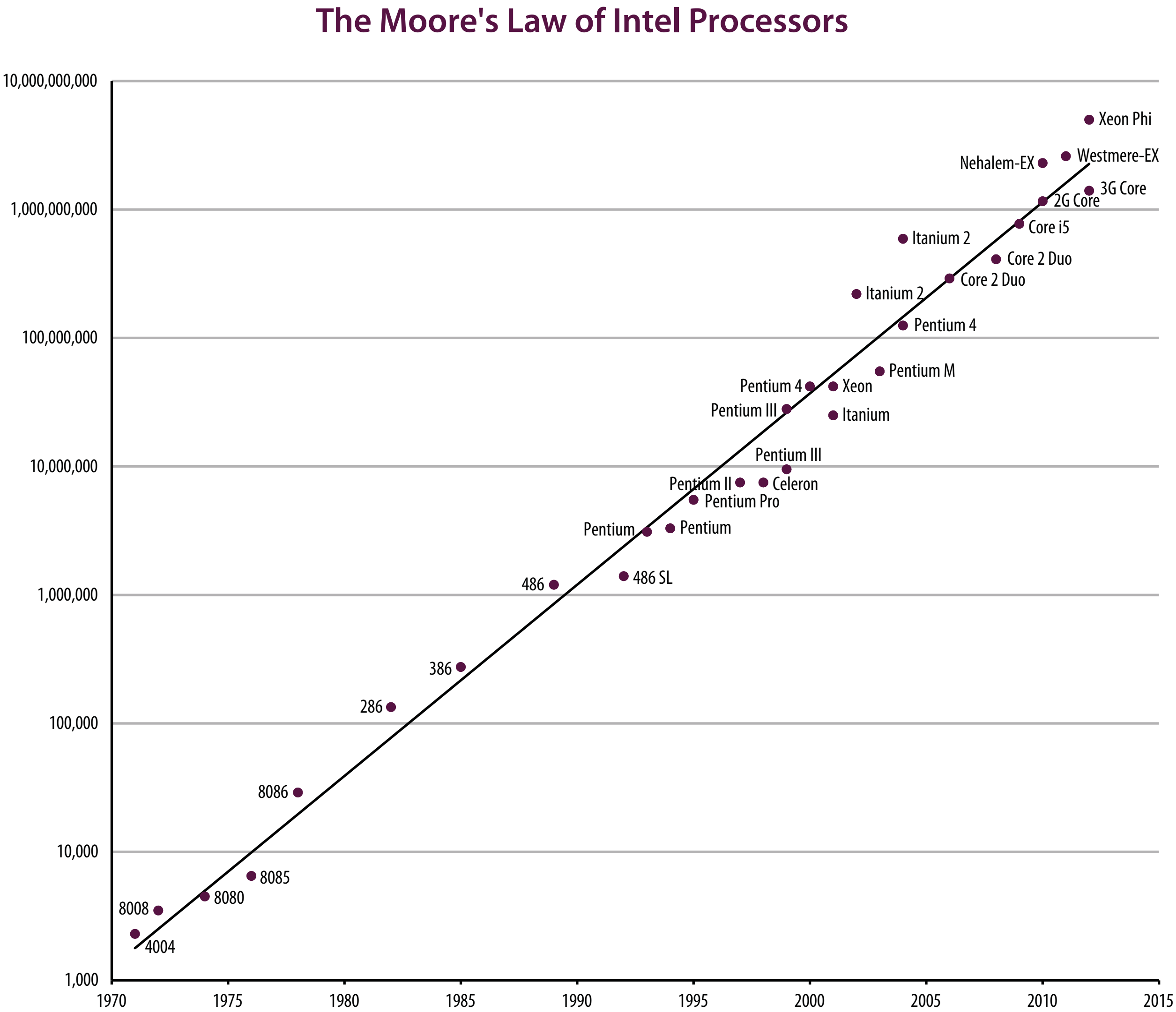


FIGURE 2. Intel has kept up with Moore's Law—the number of transistors on their processors have doubled approximately every two years.

References

[1] Moore GE. "Cramming more components onto integrated circuits." *Electronics.* 1965;38(8):114–117. doi: 10.1109/JPROC.1998.658762.

[2] Intel Corporation. "The story of the Intel 4004." Available at: http://www.intel.com/content/www/us/en/history/museum-story-of-intel-4004.html.

[3] Knight M, Glass N. "Are computer chips on the verge of a quantum leap?" *CNN.com.* 02 Sep 2013. Available at: http://www.cnn.com/2013/09/02/tech/innovation/are-computer-chips-verge-quantum/.

[4] Bohr M, Mistry K [Intel Corporation]. "Intel's revolutionary 22 nm transistor technology." 2011. Available at: http://www.intel.com/content/www/us/en/silicon-innovations/revolutionary-22nm-transistor-technology-presentation.html.

SPINOUTS

News from the Technology Transfer Program

Searching the future enterprise

NSA researcher invents system for efficient "collision-free hashing of near-match inputs"

As data services continue to expand, the problems involved in Big Data searches will increase. An invention by an NSA researcher offers a way to reduce the problem of "data collision" in enterprise-level systems.

People who manage massive databases know it's important to ensure that data relationships stay clear. When pieces of data exist in many-to-one relationships—that is, when multiple inputs map to the same output—it's possible for data to collide. Just imagine the identity problems if, say, 10 people shared the same social security number or driver's license number.

Hash functions, which map long inputs to short outputs, are important tools for efficient searching of Big Data. Unfortunately, collisions are inherent in any hash function that involves more inputs than outputs. It is possible, however, to reduce how often collisions occur. Hashes such as the message-digest algorithm (MD5) or the secure-hash algorithm (SHA) family are designed to make it infeasible to find any input that corresponds to a given output . . . but such hashes operate very slowly.

Fortunately, an NSA researcher has invented a system that could significantly reduce time spent on hashing calculations for large data sets. US patents 8,363,825 and 8,355,501 have been granted for a "Device for and method of collision-free hashing for near-match inputs." The system produces more efficient hashing calculations, which in turn enable faster hashing and data retrieval. This would benefit applications such as DNA sequencing, by enabling fast searches while guaranteeing that near matches will not collide; this would make it possible, for example, to detect single nucleotide polymorphisms (i.e., genetic variation in a DNA sequence that occurs when a single nucleotide in a genome is altered).

The invention would also be useful for applications such as the construction of Bloom filters, which record whether a searcher has already viewed a particular data record. Bloom filters, in turn, can be used for data retrieval in the cloud (or elsewhere) and can be combined with homomorphic encryption for certain applications in private information retrieval.

The collision-free hashing technology is now developed, and a software demonstration is also available. To arrange a demonstration, please contact the Technology Transfer Program at tech_transfer@nsa.gov or 1-866-680-4539.