Tip Sheet For
2019 IEEE International Electron Devices Meeting (IEDM)

For Immediate Release

The annual IEDM conference (www.ieee-iedm.org), sponsored by the IEEE Electron Devices Society, is the world’s largest, most influential forum for the unveiling of breakthroughs in transistors and related micro/nanoelectronics devices. This Tip Sheet spotlights some of this year’s most newsworthy sessions and papers. The 65th annual IEDM will be held December 7-11, 2019 in San Francisco. (See pp. 9-11 for definitions of acronyms and technical terms.)

A) IEDM Late-News Papers

- **TSMC to Unveil a Leading-Edge 5nm CMOS Technology Platform**: TSMC researchers will describe a 5nm CMOS process optimized for both mobile and high-performance computing. It offers nearly twice the logic density (1.84%) and a 15% speed gain or 30% power reduction over the company’s 7nm process. It incorporates extensive use of EUV lithography to replace immersion lithography at key points in the manufacturing process. As a result, the total mask count is reduced vs. the 7nm technology. TSMC’s 5nm platform also features high channel-mobility FinFETs and high-density SRAM cells. The SRAM can be optimized for low-power or high-performance applications, and the researchers say the high-density version (0.021µm²) is the highest-density SRAM ever reported. In a test circuit, a PAM4 transmitter (used in high-speed data communications) built with the 5nm CMOS process demonstrated speeds of 130 Gb/s with 0.96pJ/bit energy efficiency. The researchers say high-volume production is targeted for 1H20. (Paper #36.7, “5nm CMOS Production Technology Platform Featuring Full-Fledged EUV and High-Mobility Channel FinFETs with Densest 0.021µm² SRAM Cells for Mobile SoC and High-Performance Computing Applications,” G. Yeap et al., TSMC)

- **Intel Says Heterogeneous 3D Integration Can Drive Scaling**: CMOS technology requires both NMOS and PMOS devices, but the performance of PMOS lags NMOS, a mismatch which must be addressed in order to wring every last bit of performance and energy efficiency from future chips. One way to do that is to build PMOS devices with higher-mobility channels than their NMOS counterparts, but because these are built from materials other than silicon (Si) which
require different processing, it is challenging to build one type without damaging the other. Intel researchers got around this with a 3D sequential stacking architecture. They first built Si FinFET NMOS transistors on a silicon wafer. On a separate Si wafer they fabricated a single-crystalline Ge film for use as a buffer layer. They flipped the second wafer, bonded it to the first, annealed them both to produce a void-free interface, cleaved the second wafer away except for the Ge layer, and then built gate-all-around (GAA) Ge-channel PMOS devices on top of it. There was no performance degradation in the underlying NMOS devices, and in an inverter test circuit the PMOS devices demonstrated the best $I_{on}$-$I_{off}$ performance ever reported for Ge-channel PMOS transistors ($I_{on}$=497 µA/µm and $I_{off}$=8nA/µm at 0.5V). The researchers say these results show that heterogeneous 3D integration is promising for CMOS logic in highly scaled technology nodes. (Paper #29.7, “300mm Heterogeneous 3D Integration of Record Performance Layer Transfer Germanium PMOS with Silicon NMOS for Low-Power, High-Performance Logic Applications,” W. Rachmady et al., Intel.)

- **Versatile 22nm STT-MRAM Technology:** Many electronics applications require fast non-volatile memory (NVM), but embedded flash, the current dominant technology, is becoming too complex and expensive to scale much beyond 28nm. A type of embedded NVM known as STT-MRAM has received a great deal of attention. STT-MRAM uses magnetic tunnel junctions (MTJs) to store data in magnetic fields rather than as electric charge, but this ability decreases as temperature increases. That makes STT-MRAM both challenging to build – it is fabricated in a chip’s interconnect and must survive high-temperature solder reflow – and also to use in applications such as automotive, where thermal specifications are demanding and the ability to resist outside magnetic fields is critical. TSMC will describe a versatile 22nm STT-MRAM technology that operates over a temperature range of -40ºC to 150ºC and retains data through six solder reflow cycles. It demonstrated a 10-year magnetic field immunity of >1100 Oe at 25ºC at a 1ppm error rate, and <1ppm when in a shield-in-package configuration. The researchers say that by trading off some of the reflow capability and using smaller MTJs, even higher performance can be achieved (e.g., 6ns read times/30ns write times), making them appealing for artificial intelligence inference engines. (Paper #2.7, “22nm STT-MRAM for Reflow and Automotive Uses with High Yield, Reliability and Magnetic Immunity and with Performance and Shielding Options,” W. Gallagher et al., TSMC)

**B) Focus Sessions**

The 2019 IEDM will feature special Focus Sessions on the following topics:

- **Emerging AI Hardware Technologies** – Increased semiconductor capabilities open up the possibility to further increase computers’ abilities to recognize signals and patterns, to make decisions based on that input, and then to act on those decisions on their own. Among other topics, this Focus Session will discuss various in-memory processing approaches for faster, more energy-efficient AI processing, analog computing techniques, the use of different device types for network “neurons,” and a photonics-based approach for faster neural networks. (Session #22)
  - Design Considerations for Efficient Deep Neural Networks on Processing-in-Memory Accelerators, Vivian Sze, MIT
  - Towards 10000TOPS/W DNN Inference with Analog in-Memory Computing – A Circuit Blueprint, Device Options and Requirements, Stefan Cosemans, imec
  - The Marriage of Training and Inference for Scaled Deep Learning Analog Hardware, Tayfun Gokmen, IBM
  - Can In-Memory/Analog Accelerators be a Silver Bullet for Energy-Efficient Inference?, Jun Deguchi, Kioxia (formerly Toshiba Memory)
Human Machine Interface – Electronic systems in different forms, such as wearables (e.g., fitness trackers), humanoid robots and AR/VR goggles, are making inroads in many areas for applications such as communications, health, entertainment, industrial productivity, and others. This Focus Session will survey and explore key issues related to the interface of these bioelectronic systems with humans. It will address challenges related to computer vision and image sensing capabilities, as well as the development of haptic sensors, and also bioelectronic interfaces with humans for the purpose of medical diagnostics and healthcare. (Session #10)

Quantum Computing Infrastructure – Quantum computing has emerged as a possible candidate to address specific computational challenges not reachable by today’s classical machines, because it exploits the laws of quantum physics and may make much more powerful and/or specialized computers possible. Future scaling of quantum computing beyond a few tens of qubits will also entail challenges for the classical electronics and systems surrounding the qubits themselves. This Focus Session will explore system-level challenges of quantum computing, such as relevant semiconductor-related fabrication and interconnect issues, and control electronics, and will brainstorm R&D directions for new materials, devices, circuits, manufacturing approaches and benchmarks for the scalable integration of a large number of qubits with CMOS technology, operating at cryogenic temperatures. (Session #31)

Reliability and Security in Circuits & Systems – Reliability is a major design concern for computing systems, given the many inter-relationships between devices and circuits, plus the increasing need to detect/resist malware at the circuit level. This Focus Session will explore topics ranging from high-level discussions of major reliability issues to specific approaches to address key challenges. (Session #13)
C) Advances in Core CMOS Technology

- **Ferroelectric Gate Dielectric Boosts Drive Current**: In transistors, the insulator that prevents current leakage from the gate is called a gate dielectric. A dielectric’s resistance to current flow, or its relative permittivity, is “k”, and a higher k leads to a higher insulating capability under the same equivalent oxide thickness. As devices scale smaller, gate dielectrics become thinner, and it has been challenging to find gate dielectric materials with a suitable k-value that are compatible with typical chip-fabrication processes. The industry is now at a point where there is almost no space left for further reducing the physical thickness of gate dielectrics without incurring unacceptable levels of gate leakage current. A Notre Dame-led team will propose making use of the ferroelectric properties of a hafnium zirconium oxide gate dielectric (Hf_{1-x}Zr_xO_{2}). Their theoretical modeling and experimental data showed that by carefully engineering the composition of the film near its morphotropic phase boundary (an electrical phase-transition region in ferroelectric materials), it is possible to achieve a k-value of 38, which can improve transistor electrostatics and boost drive current by 13%. (Paper #7.4, “Equivalent Oxide Thickness (EOT) Scaling with Hafnium Zirconium Oxide High-k Dielectric Near Morphotropic Phase Boundary,” K. Ni et al., Univ. Notre Dame/Purdue Univ./Kurt J. Lester Co.)

- **3D Stacking at Low Temperature**: A team led by Taiwan Semiconductor Research Institute will describe the monolithic 3D fabrication of GAA transistors at less than 400ºC. The relatively low processing temperature was key to their ability to build and stack transistors without damaging existing devices. The stackable structure had a 5nm-wide, 34nm-high rectangular polycrystalline nanowire channel. The GAA devices demonstrated a record-high I_{on}/I_{off} ratio (10^8) and low I_{off} current (pFETs<10^{-2}nA/µm). Key enablers were plasma-assisted atomic layer etching for channel sidewall patterning, and plasma-immersion ion implantation and far-infrared laser pulsing for optimized activation. The researchers demonstrated the advantages of this stacked structure in an SRAM circuit for compute-in-memory applications. (Paper #3.3, “Monolithic SRAM-CIM Macro Fabricated with Stackable Gate-All-Around MOSFETs,” F.K. Hsueh et al., Taiwan Semiconductor Research Institute/National Chiao Tung Univ/UC-Berkeley)

- **Vertical Devices for Ultra-Scaled Circuits**: imec researchers will describe p- and n-type vertical GAA nanowire and nanosheet transistors with SiGe/Si pillars and self-aligned spacers, which could lead to ultra-scaled circuits. The researchers carried out comprehensive studies on the impacts of device doping and physical dimensions on transistor performance, variability, noise and reliability, for junctionless devices vs. inversion-mode devices built with a replacement metal gate process flow. To build the self-aligned spacers that align the gate with the source/drain, the researchers made use of the different oxide growth rates of Si and Ge. They propose that when used as the selector for an STT-MRAM, these vertical devices would allow for...
a substantial area reduction (64%) vs. FinFETs, along with substantial reductions in read/write energy and latency. (Paper #11.1, “Vertical Nanowire and Nanosheet FETs: Device Features, Novel Schemes for Improved Process Control and Enhanced Mobility, Potential for Faster & More Energy-Efficient Circuits,” A. Veloso et al., imec)

- **Electrical Characterization of MoS2**: Molybdenum disulfide (MoS2) is a so-called 2D material because it has length and width but is only one atomic layer thick. Its electrical properties may make it suitable for use in future extremely scaled devices. A team led by imec studied how the size and spacing of electrical contacts to MoS2 affect its performance. They built devices with the smallest contact pitch reported for 2D MOSFETs (42nm) and the smallest contact length (13nm). These devices demonstrated a high on-current for MoS2 with a 50nm SiO2 gate oxide ($I_{on}=250\mu\text{A}/\mu\text{m}$). When a much thinner 4nm HfO2 gate oxide was used, they achieved an excellent subthreshold slope (80mV/dec). (Paper #23.2, “Ultra-Scaled MOCVD MoS2 MOSFETs with 42nm Contact Pitch and 250\mu\text{A}/\mu\text{m Drain Current},” Q. Smets et al., imec/KU Leuven)

**D) Neural Networks/Artificial Intelligence/Neuromorphic Computing**

- **Fully Integrated Spiking Neural Network**: CEA-Leti will describe the first complete integration of a spiking neural network (SNN) on a chip. Spiking neural networks are energy-efficient and attempt to mimic biological neural functioning by classifying information based on ‘spikes,’ or discrete events occurring at a point in time vs. continuous values. CEA-Leti built a 130nm CMOS test chip with analog neurons and resistive-RAM-based (RRAM) synapses, monolithically integrated on top of CMOS devices. It demonstrated a classification accuracy of 84% on the MNIST database of handwritten digits (used to train image-processing systems), with 5x lower energy dissipation at the synapse and neuron level (3.6 pJ) vs. other chips that use formal programming methods for image classification. The researchers say that moving to the 28nm node from 130nm could result in a 10x energy reduction and a 30x density gain, and that using RRAM to build multiple-level memory cells vs. the single-level cells in the test chip can further reduce synaptic density by 4x. (Paper 14.3, “Fully Integrated Spiking Neural Network with Analog Neurons and RRAM Synapses,” A. Valentian et al., CEA-Leti)

- **Optimized Flash Memory for DNNs**: Extensive multiply-and-accumulate mathematical computations lie at the heart of deep-learning neural networks (DNNs). It is energy-intensive to move data back and forth between processor and memory to carry them out, though, and various “compute-in-memory” (CIM) approaches to bring the two closer together are being explored. Macronix researchers will discuss how non-volatile single-level cell (SLC) 3D NAND flash memory has density and cost advantages vs. RRAM for high-density DNNs (>100Mb). They will describe how they optimized it for use in a CIM accelerator for DNNs. They built a 16-layer SLC single-gate, vertical-channel 64Gb chip with ultra-low on-current ($I_{on}=2n\text{A}$) and off-current ($I_{off}$ much <1pA). The high on/off ratio provides a large bandwidth, enabling >10,000 memory cells to be summed in parallel, resulting in high energy-efficiency (TOPS/W~40) along with high accuracy and error-tolerance. (Paper 38.1, “Optimal Design Methods to Transform 3D NAND Flash into a High-Density, High-Bandwidth and Low-Power Nonvolatile Computing-in-Memory (nvCIM) Accelerator for Deep-Learning Neural Networks,” H.T. Lue et al., Macronix)

**E) Advanced Technology Platforms in Wide-Ranging Areas**
• **STT-MRAMs for L4 Cache**: Intel researchers will present a step forward for STT-MRAM technology, in a paper that details array-level MTJ process, performance and reliability requirements for the use of STT-MRAMs in L4 cache operations. They built 2MB arrays of scaled MTJ devices capable of meeting L4 cache specifications across all operating temperatures. The chips demonstrated a 20ns write time, 4ns read time, endurance of $10^{12}$ cycles, and memory retention of one second at elevated temperature (110ºC). Bit error rates were low enough that they could be detected and corrected with error-correcting code (ECC) techniques. (Paper #2.4, “2MB Array-Level Demonstration of STT-MRAM Process and Performance Towards L4 Cache Applications,” J.G. Alzate et al., Intel)

• **A 400 Gbits/s Silicon Photonics Platform**: Data is shuttled around within data centers via optical fiber, using silicon photonic (light-based) interfaces that operate at 100 Gbits/s. But cloud data center traffic is growing at nearly 30% per year and there soon will be a need to increase the data rates. An STMicroelectronics-led team will describe a new silicon photonics technology platform built on 300mm SOI wafers, yielding devices that operate at 400Gbit/s (each device has 4 channels, each of which operates at 100Gbit/s, for a total of 400Gbit/s). Optical coupling and polarization management are key requirements, and their devices incorporate a 60 GHz high-speed photodiode and a high-efficiency, high-speed phase modulator. They also built devices with a supplementary SiN waveguide layer for higher coupling efficiency, to meet evolving data-transmission requirements. The researchers say the photonics platform has the potential to meet the requirements of applications other than data centers, too, such as automotive. (Paper #33.1, “A Silicon Photonics Technology for 400Gbit/s Applications,” F. Boeuf et al., STMicroelectronics/Univ. di Pavia)

F) **Device Technology Advancements for 5G and Beyond**

• **3D Heterogeneous Integration of GaN and Si**: Intel researchers will present a second paper on 3D heterogeneous integration in addition to the late-news presentation. This one will describe how low-leakage, high-k dielectric enhancement mode GaN NMOS and Si PMOS transistors were built monolithically on a 300mm high-resistivity Si substrate. The GaN devices were built using MOCVD epitaxy, and the Si devices were built on top of them by means of a layer transfer process. The goal is to combine GaN’s high-frequency/-temperature/-power attributes with silicon CMOS circuitry’s digital signal processing, logic, memory and analog capabilities, to create compact devices for next-generation solutions for power delivery, RF and system-on-chip (SoC) applications. The researchers say both device types demonstrated excellent performance across a range of electrical specifications. For example, the GaN NMOS devices showed a low off-current (100pA/µm) and excellent RF performance/efficiency at both mmWave and sub-7GHz frequencies, while the Si PMOS devices demonstrated a high drive current (0.85mA/µm) and a low off-current (150pA/µm). (Paper #17.3, “3D Heterogeneous Integration of High-Performance High-k Metal Gate GaN NMOS and Si PMOS Transistors on 300mm High-Resistivity Si Substrate for Energy-Efficient and Compact Power Delivery, RF (5G and Beyond) and SoC Applications,” H. W. Then et al., Intel)

• **III-V HBTs on Si**: III-V materials offer higher electron mobilities than silicon, and HBTs made from them are very fast transistors often used for RF and other high-frequency applications. A key goal is to build them on 300mm silicon wafers instead of other substrates, to take advantage of silicon’s lower manufacturing costs. A team led by imec will describe how they used a unique nano-ridge engineering technique to build GaAs/InGaP HBTs on a 300mm silicon substrate. Key
to the success of the technique is an aspect ratio trapping mechanism by which defects are confined to narrow trenches. The HBTs demonstrated excellent electrical performance, with a DC current gain of ~112 and a breakdown voltage of 10V. The researchers say the technique can also be used to build InGaAs devices on 300mm Si substrates. (Paper #9.1, “First Demonstration of III-V HBTs on 300mm Si Substrates Using Nano-Ridge Engineering,” A. Vais et al., imec/KU Leuven)

- **RF & Mixed-Signal Reliability Challenges**: With the advent of increased wireless connectivity, there is a much greater focus on the reliability of radio-frequency (RF) and mixed-signal circuits. In line with this year’s theme, “Innovative Devices for an Era of Connected Intelligence,” Texas Instruments researchers will present a deep review and summary of the reliability challenges encountered during the design of RF and mixed-signal circuits. (Paper #13.1, “Challenges in Radio Frequency and Mixed-Signal Circuit Reliability,” V. Reddy et al., Texas Instruments)

**G) Power Devices**

- **Outstanding GaN-on-Si Performance**: Lateral GaN-on-Si devices are promising for advanced power ICs because they combine high device performance, low cost, and can make physically smaller systems possible. While GaN HEMTs are commercially available, there is still plenty of room to improve their performance up to the full potential of GaN, and further advances require a major advance in on-resistance while still maintaining high breakdown voltage. A team led by EPFL will report on novel lateral multi-channel AlGaN/GaN power devices with high breakdown voltage (1230 V) and low on-resistance (2.8 Ω-mm), resulting in an excellent figure-of-merit of 3.2 GW/cm². The researchers call these devices MOSHEMTs, and they have a multi-channel FinFET-like architecture with slanted tri-gates whose structure can be modified to “tune” device performance. The on-resistance is some 5x lower than previously seen at this breakdown voltage with single-channel devices. Until now it has proven difficult to control multiple parallel channels with a typical gate architecture, but the tri-gate structure has proven to be effective in doing so, and it lends itself to building normally-off – and thus more fail-safe – devices. (Paper 4.1, “1200 V Multi-Channel Power Devices with 2.8 Ω-mm ON-Resistance,” J. Ma et al., EPFL/Enkris Semiconductor)

- **Monolithically Integrated GaN-on-SOI Platform Technology**: A team led by imec will discuss monolithically integrated GaN-on-SOI ICs for power conversion, which they used to build a complete buck converter that demonstrated 200 V on-chip power conversion at 1 MHz. An SOI (silicon-on-insulator) substrate was used because it eliminates the back-gating effect which wastes power. It also suppresses parasitics, provides effective electrical isolation and reduces the area required. The team comprehensively investigated the technology from multiple vantage points: substrate, buffer, isolation, device, co-integration and circuit. Different components were successfully integrated monolithically, including a HEMT, Schottky barrier diode, MIM capacitor, 2DEG resistor, and resistor-transistor logic. (Paper 4.4, “GaN-on-SOI: Monolithically Integrated All-GaN ICs for Power Conversion,” X. Li et al., imec/KU Leuven)

**H) Noteworthy Papers on Diverse Topics**

- **Giving Forceps a Sense of Touch**: Our internal organs are slippery because they’re covered with blood and other body fluids, so grasping and pulling them with forceps can be challenging.
Although contact-force sensors have been placed on the tips of forceps used in laparoscopic and robotic surgeries, there currently is no way to know if they are slipping, other than visually via a monitor, which has limited usefulness. A Kagawa University team will describe a highly sensitive slip-sensing imager (sub-mm resolution) and novel algorithm that can, in effect, give forceps a sense of touch. The idea is to use the device to visualize the spatial distribution of the grasping force across the organ’s surface. The center of that distributed load is calculated, and as the forceps are moved the algorithm relates any corresponding movements of the load center to slippage. Built on an SOI wafer, the device’s force-sensor pixels consist of a 20µm–thick piezoelectric silicon diaphragm (400µm diameter) with a center contact, and with a force detection circuit integrated on the diaphragm. The diaphragm acts as a strain gauge as it flexes due to varying grasping force. (Paper #18.2, “Highly Sensitive Slip Sensing Imager for Forceps Grippers Used Under Low Friction Condition,” K. Ando et al., Kagawa Univ./Takamatsu Red Cross Hospital)

- **High Energy-Density Thin-Film Battery**: There has been great progress in miniaturizing electronics but the miniaturization of power sources hasn’t kept pace. Although integrated electrochemical capacitors offer high power density, high frequency response and novel form factors, their low energy densities are of limited value for MEMS and autonomous device applications that require long periods between charging. CEA-Leti researchers will discuss a thin-film battery with the highest areal energy density yet reported (890 µAh/cm²) and high power density (450 µAh/cm²). Built on silicon wafers using UV photolithography and etching for the successive deposition and patterning of each layer, the thin-film battery integrates a 20µm-thick LiCoO₂ cathode in a Li-free anode configuration. It showed good cycling behavior over 100 cycles, and the fact it was built using a wafer-level process opens up the possibility to tightly integrate this battery technology with future electronic devices. (Paper 26.1, “Millimeter-Scale Thin-Film Batteries for Integrated High Energy-Density Storage,” S. Oukassi et al., CEA-Leti)

- **Design Framework for Quantum Computing**: Design software such as TCAD is used to produce highly accurate models of semiconductor devices and their operation, but no analogous tools exist to model qubits, the basis of quantum computing, because the field is so new and complex. If these design tools did exist, the development of quantum computers could take place much more quickly. A team led by imec has taken a step to create such a software framework, and will describe how they used multiphysics simulation methods to develop a comprehensive design methodology for qubits built in silicon. They modeled device electrostatics, stress, micromagnetics, band structure and spin dynamics. Based on the results of these studies, they say that single-electron qubits in quantum dots can be induced and optimized in silicon MOSFETs with thin (<20nm) gate oxides. The researchers will discuss critical aspects of their methodology, the parameters they modeled, and next steps. (Paper #39.5, “Multiphysics Simulation & Design of Silicon Quantum Dot Qubit Devices,” F.A. Mohiyaddin et al., imec/KU Leuven/ETH Zurich)

- **Monolithic Impedance Sensor for Fingerprint Imaging**: Researchers led by Cornell will discuss the monolithic integration of a piezoelectric AlN resonator into a CMOS-controlled, GHz ultrasonic impedance sensor/imager. The device measures changes in surface properties such as surface oxidation, materials, liquid viscosity and others, and is meant for use in wearable, IoT and smartphone systems to detect fingerprints with high resolution, determine tissue states, and for other applications. This is the first time monolithic fabrication has been successfully demonstrated, and it led to small, power-efficient GHz sensing arrays with improved
performance vs. the standard two-chip heterogeneous integration approach, thanks to less parasitic coupling and a higher signal-to-noise ratio. (Paper #34.3, “Monolithic 180nm CMOS-Controlled GHz Ultrasonic Impedance Sensing and Imaging,” M. Abdelmejeed et al., Cornell Univ./A-STAR)

- **Ultra-Low-Power Cryptographic Solution:** The spread of networked mobile devices and smart gadgets in the IoT landscape has created an urgent need to protect them with lightweight and low-power cryptographic solutions. A physically unclonable function (PUF) is a hardware-intrinsic security primitive, or basic programming element. UC Santa Barbara researchers will discuss an ultra-low-power PUF that operates on the varying electrical resistances and current leakages that arise from intrinsic process variations in ReRAM crossbar arrays. The team built 4K-ReRAM passive crossbar circuit arrays fabricated with a CMOS-compatible process suitable for back-end-of-the-line (BEOL) integration. The arrays allow for an extremely large number (~10^25) of challenge-response pairs (a common cryptographic protocol), as well as 4x better density vs. other ReRAM architectures plus a ~100x improvement in power efficiency and more robust security metrics. (Paper #30.1, “Ultra-Low Power Physically Unclonable Function with Nonlinear Fixed-Resistance Crossbar Circuits,” M.R. Mahmoodi et al., UC-Santa Barbara)

- **High-Definition Infrared Image Sensor:** Demand for imaging in the short-wavelength infrared range (SWIR, or 1,000-2,000nm wavelengths) has been increasing for industrial, science, medical, agricultural and security purposes. InGaAs has been used to build SWIR sensors because it can absorb light in this range that silicon cannot. With conventional back-illuminated InGaAs sensors, each pixel of a photodiode array is connected to a readout circuit on a silicon wafer by means of a microbump. But it’s difficult to scale these bumps, and so creating fine-pitch pixel arrays for greater image definition is difficult. A Sony team will describe an architecture in which each pixel in an InGaAs/InP photodiode array is connected to the readout circuit not with microbumps, but by means of copper-to-copper bonding, resulting in a much tighter pitch. They used the technique to build a prototype 1280 x 1024-pixel array with a 5µm pitch. Also, thinning of the InP layer and process optimization yielded a sensor that demonstrated high sensitivity and low dark current, respectively. The researchers say this work paves the way for high-definition SWIR imaging. (Paper #16.7, “High-Definition Visible-SWIR InGaAs Image Sensor Using Cu-Cu Bonding of III-V to Silicon Wafer,” S. Manda et al., Sony)

Here are definitions of some important technical terms you may find useful:

- **Back-End/BEOL** and **Front-End/FEOL** -- In integrated circuit manufacturing, transistors and other active devices are built first (at the front end of the manufacturing line or FEOL), while the interconnect, or the wiring, is built afterward, at the “back end” of the manufacturing line (BEOL). The BEOL imposes a much lower thermal budget than the FEOL which includes doping activation and gate stack anneals.
- **Carbon nanotube (CNT)** – A cylinder made of carbon atoms, measured in nanometers (nm). Depending on its chirality it may possess metallic or semiconducting properties.
- **CMOS/MOS/MOSFET/FET**-- Most transistors today are FETs, or field-effect transistors. CMOS (complementary metal oxide semiconductor) technology requires FET devices of both polarities (n-FETs and p-FETs) which turn on at the application of a positive or negative gate bias, respectively. Generically they are called MOSFETs, or sometimes MOS transistors.
- **Compound/III-V Semiconductors** -- Most semiconductors are silicon-based, but researchers continue to investigate other semiconducting materials with higher electron or hole mobilities because they can be used to make faster devices. The tradeoff is that the materials are harder to work with than silicon. Compound semiconductors are
made of two or more elements (e.g. GaAs, InP, GaN, etc.) that are generally found in groups III and V of the periodic table of the elements.

- **DHBT** – Double heterojunction bipolar transistor
- **DIBL** – Drain-induced barrier lowering is a parasitic short-channel effect in MOSFETs that influences threshold voltage.
- **Electromigration** – A serious reliability issue. At tiny dimensions some materials tend to physically move when current flows through them, leading to voids, gaps or outright breaks in what should be a uniform material.
- **Electron mobility** – A measure of the velocity of electrons in a semiconductor. The corresponding measure for positive charge, or holes, is termed hole mobility.
- **FinFET** – A transistor whose shape resembles a fin, usually with multiple gates surrounding it for better on/off control.
- **f_T** – a measure of transistor speed known as unity current gain cutoff frequency, above which the device loses its amplifying capability.
- **f_{max}** – a measure of transistor speed denoting its maximum oscillation frequency
- **Gate-all-Around (GAA) Device** – A transistor with multiple gates surrounding the channel on all sides.
- **HEMT** – A HEMT, or high-electron-mobility transistor, is a field-effect transistor with a channel built from a sandwich of two materials with different energy band gaps, instead of a channel built from a semiconductor and separated from the gate electrode by an insulating oxide, as is the case with MOSFETs. HEMTs operate at higher frequencies than ordinary transistors, up to millimeter-wave frequencies, and are used in high-frequency products such as cell phones, satellite television receivers, voltage converters, and radar equipment.
- **High-\(k\) Dielectrics/Metal Gates (HKMG)** – A dielectric is an insulator, and “\(k\)” is relative permittivity, a measure of how high capacitance will be under the same physical thickness. A higher numerical value for “\(k\)” leads to a better insulating performance without increased physical thickness of the insulating layer. For the gate of a transistor, which turns it on and off, a high-\(k\) dielectric is critical because if current leaks through the gate, the transistor won't work properly. In the past gates were usually made of polysilicon with an oxide dielectric, but these are not scalable to meet the needs of the latest technology generations. Other good dielectrics exist but they aren't compatible with polysilicon gates. The latest generations of MOSFETs use metal gates with high-\(k\) dielectrics to enable continued scaling and higher device performance.
- **III-V** – see Compound/III-V Semiconductors
- **Integrated Circuit** – A tiny electrical circuit built on a semiconducting substrate.
- **Internet of Things (IoT)** – The idea of giving everyday objects network connectivity, allowing them to send and receive data.
- **Lithography/Photolithography/EUV Lithography** – Transistors are built on chips using photolithography, a process by which light is shone through a patterned photomask onto a chip’s surface. The light transfers a copy of the image on the mask to a light-sensitive layer on the chip’s surface, in much the same way the image on a photographic negative is transferred onto photo paper. The chip’s surface is then engraved with chemical etches and other treatments in order to build the patterned devices and circuits. Right now various tricks and tweaks enable us to build transistors smaller than the wavelength of light used to pattern them. But for future technology generations of ultra-small transistors, the shorter wavelengths of extreme ultraviolet light (EUV) may be needed. Researchers are working to develop cost-effective, high-throughput EUV systems.
- **Low-\(k\) Dielectrics/Interconnect** – Interconnect refers to the copper lines that connect devices on a chip. The tiny widths and close proximity of adjacent lines introduce resistance and capacitance delays that can hinder chip performance. Here a low-\(k\) dielectric is needed, to insulate the copper lines while minimizing capacitance increase, but these materials are fragile and pose many challenges.
- **MEMS/NEMS** – A micro-electromechanical system/nano-electromechanical system, containing micrometer-scale moving parts (the former) or nanometer-scale moving parts (the latter).
- **MRAM** – Magnetoresistive Random Access Memory, a type of non-volatile memory that stores data not as an electrical charge, but as a magnetic state. See STT-MRAM.
- **Multi-chip Modules** – An electronic package containing multiple ICs, other semiconductors and/or other chip modules.
- **Moore’s Law** – An observation made in 1965 by Intel’s Gordon Moore, who said that the number of transistors which can be placed on an integrated circuit doubles about every two years. Subsequent events have proved him correct but it’s getting harder to do.
• **Negative capacitance** – A mathematical measure of the momentary change in capacitance of a ferroelectric device during switching, which can aid in the design of more energy-efficient ferroelectrics. A ferroelectric material switches its polarization when a critical voltage is applied. In some cases that voltage causes a large electrical charge to build suddenly on the material’s surface. That surface charge may quickly grow to exceed the magnitude of the charge already being supplied to the device from an external power supply. If at that moment a resistance is introduced between the device’s electrodes and the charge from the power supply is reduced, then the voltage would decrease even as the overall level of charge in the ferroelectric material is still increasing because of the surface accumulation. Because capacitance is charge divided by voltage, it would have a negative value at that moment.

• **N-FET/P-FET or n-channel/p-channel** -- All MOSFETs come in two varieties that work together in a complementary fashion, N and P. These act like switches which turn on at a positive and a negative gate bias, respectively.

• **Phase-Change Memory/PCM** -- Phase-change materials have crystalline and non-crystalline states that are used to represent the digits “0” or “1,” the basis of computer memories. Electrical current is used to toggle between the two states – heat from the current melts the material to change its state.

• **Scaling/Density/Integration** -- Scaling is making transistors and other circuit elements smaller so that more will fit on a chip. A denser chip has more transistors on it than one which is less dense. Integration is combining circuit elements on a chip to add more functions at less cost.

• **Self-Assembly/“Bottom Up”** -- A manufacturing method used at the nanometer scale, making use of the fact many biological systems automatically self-assemble into various molecular or other structures. Self-assembly techniques imitate these strategies by encoding a desired structure into the shape and properties of the molecules that are used. This compares to traditional lithography, where the structure is carved out of a block of matter. Self-assembly is thus referred to as a ‘bottom-up’ manufacturing technique, as compared to lithography, a “top-down” technique.

• **Semiconductor** -- A material that can be made to conduct or to block the passage of electrical signals, giving the ability to store and process data.

• **SOI** -- A silicon-on-insulator substrate, used to reduce parasitic capacitance and thereby improve performance.

• **Strained silicon & SiGe stressors** -- Silicon is said to be “strained” when its atoms are pulled farther apart or closer together than normal. Doing so alters the structure of the material and, hence, the atomic forces that govern the flow of electrons through the silicon, enabling transistors built with it to operate faster and/or at lower power. The external stressors which impart strain are materials with slightly different atomic structures than silicon. For example, a common way to strain silicon is to build a silicon layer on top of a silicon-germanium (SiGe) layer. The silicon atoms will align with the atoms of the SiGe layer, which lie farther apart.

• **SRAM** -- A type of computer memory (static random access memory) that uses six transistors to store each bit of information. It can be written to and read from very quickly.

• **STT-MRAM** -- Spin-Transfer Torque Magnetoresistive RAM, an advanced type of MRAM that offers fast write speeds, the potential for very high density, low-power operation, and long endurance.

• **Subthreshold Swing or Inverse Subthreshold Slope** -- A measure of a MOSFET transistor’s current/voltage characteristics. The steeper it is, the more abruptly the device can turn on and off.

• **Technology Generations/Nodes** -- Production of 22-nm and 14nm devices by leading companies is established and the transition to 7nm/10nm is well underway.

• **Transconductance (Gm)** -- A measure of transistor performance. Transistors with higher levels of transconductance can amplify signals more efficiently than low-transconductance devices can. Technically, in a transistor it is the ratio of change in output current to change in input voltage.

• **Transistor** -- A tiny electrical switch that makes electronic systems possible. It has no moving parts and is made from semiconductors, usually silicon. Transistors can be ganged together by the millions on chips and programmed to sense inputs, to process information and to deliver outputs.

• **TSV** – Through-silicon via, a vertical electrical connection (i.e. a “via”) that passes through a silicon wafer or die. TSVs are used to create 3D integrated circuits and packages of stacked microchips.