Technical Highlights from the 2019 Symposia on VLSI Technology & Circuits

The 2019 Symposia on VLSI Technology & Circuits is a premiere international conference that defines the pace, progress and evolution of microelectronics, scheduled from June 9-14, 2019, in Kyoto, Japan. The two Symposia feature a fully overlapping technical program that includes many joint sessions. The Symposia is preceded by Sunday Workshops on June 9 and full day Short Courses on June 10, and followed by a Friday Forum dedicated to enabling technologies for autonomous driving on June 14.

Built around the theme of **"Pushing the Limits of Semiconductors for United and Connected World,"** the Symposia program integrates advanced technology developments, innovative circuit design, and the applications that they enable as part of our global society's adoption of smart, connected devices and systems that change the way humans interact with each other.

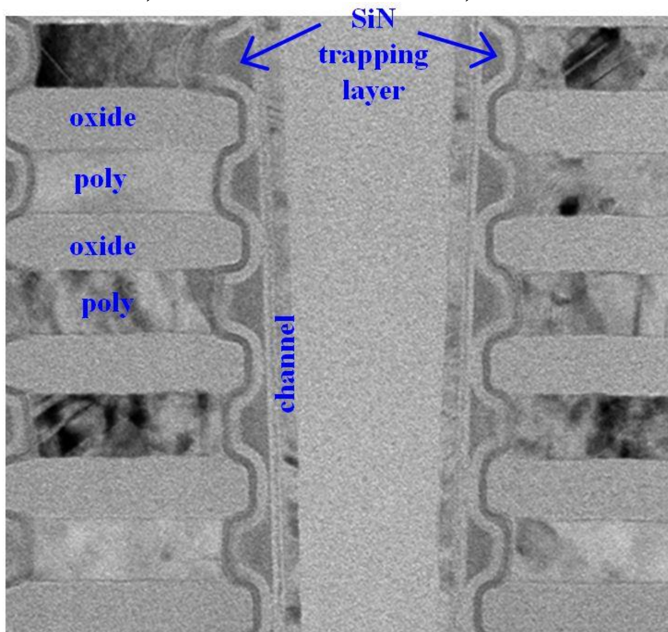
Following are some of the highlighted papers that address this theme:

ADVANCED MEMORIES

3D NAND with Confined Nitride-Trapping Layer for Robust Retention

Data retention is a key challenge especially in nitride-trapping layer type 3D NAND Flash. Macronix International Co., Ltd. will present the 3D NAND Flash with a confined nitride (SiN) trapping layer. This structure demonstrates excellent retention with only ~600 mV shift of charge loss (out of initial 7 V window) after 125 ℃ 1-week high-temp baking for a post 1K cycled device. It is far superior to the control sample without confined SiN structure.

Paper T16-2 "A Novel Confined Nitride-Trapping Layer Device for 3D NAND Flash with Robust Retention Performances,"
C.-H. Fu et al., Macronix International Co., Ltd.

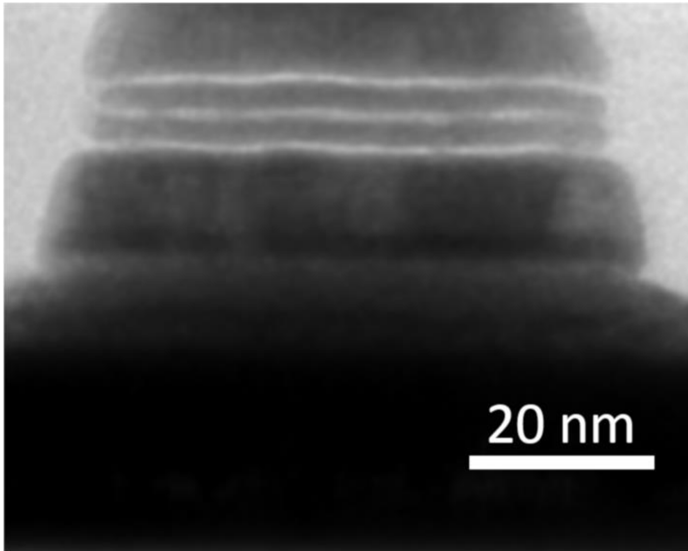


T16-2 Cross sectional TEM image of 3D NAND Flash structure with a confined SiN trapping layer.

STT-MRAM technology beyond 20nm

For the scaling of STT-MRAM beyond 20nm, Tohoku University proposes novel quad-interface magnetic tunnel junction (MTJ) technology using 300mm process based on novel low damage integration process. It is regarded as post-double-interface MTJ technology, and shows increase of both thermal stability factor delta and switching efficiency delta/IC0 by a factor of 1.5-2 compared with conventional double-interface MTJ technology.

Paper T11-4 "Novel Quad interface MTJ technology and its first demonstration with high thermal stability and switching efficiency for STT-MRAM beyond 2Xnm,"
K. Nishioka et al., Tohoku Univ.

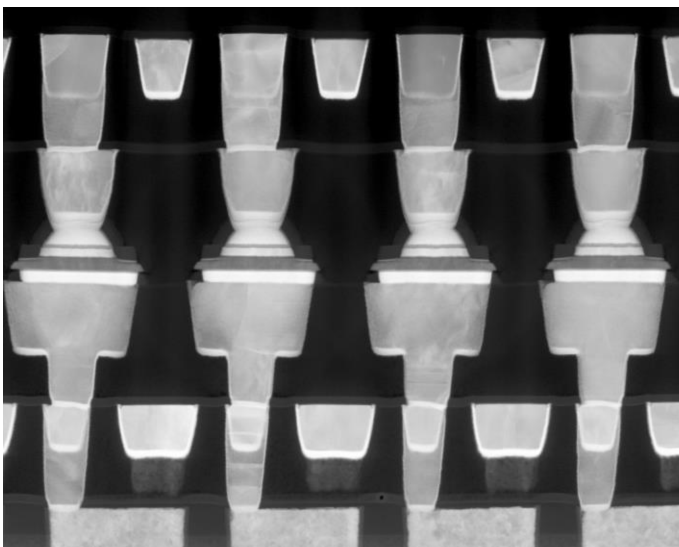


T11-4 Cross-sectional TEM image for the fabricated quad-interface MTJ.

Embedded RRAM on 22FFL FinFET technology

Embedded memory technology with logic friendly process, good retention and endurance characteristic is required in the market. RRAM is the promising candidate for this requirement. Intel Corp. will present the technology of embedded RRAM on 22FFL FinFET. 10^4 cycle endurance combined with 85°C 10-year retention is achieved on 7.2Mbit arrays.

Paper T18-1 "Non-Volatile RRAM Embedded into 22FFL FinFET Technology,"
O. Golonzka et al., Intel Corp.

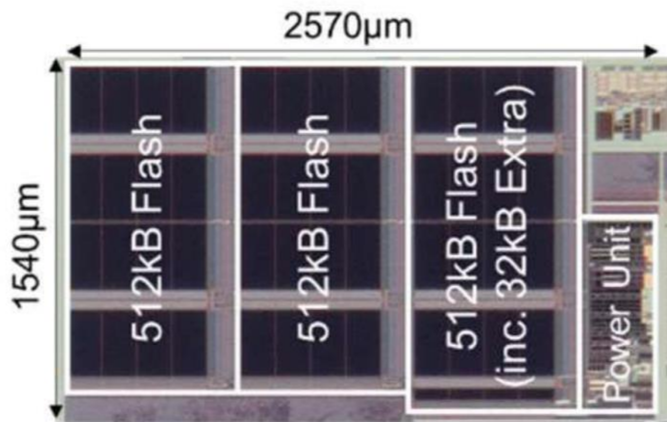


T18-1 Cross-sectional TEM of an RRAM array embedded in metal interconnect of the 22FFL logic.

Low Power Embedded Flash Memory for IoT Applications

To expand IoT application ranges, ultra-low power operation is a key requirement of many edge devices. Specifically, read energy reduction in embedded Flash (eFlash) is strongly required to enable real-time sensing with limited energy generation by energy harvesting (EH). Renesas Electronics will present a 1.5MB 2T-MONOS eFlash macro fabricated with 65nm Silicon-on-Thin-Box (SOTB) technology, using low-energy sense amplifier and data transmission circuit techniques which enhance advantages of SOTB devices. The proposed eFlash achieves 0.22pJ/bit read energy with 64MHz read access, which is low enough to utilize EH technologies as energy source.

Paper C17-1 "A 65nm Silicon-on-Thin-Box (SOTB) Embedded 2T-MONOS Flash Achieving 0.22 pJ/bit Read Energy with 64 MHz Access for IoT Applications,"
K. Matsubara, et al., Renesas Electronics Corp.



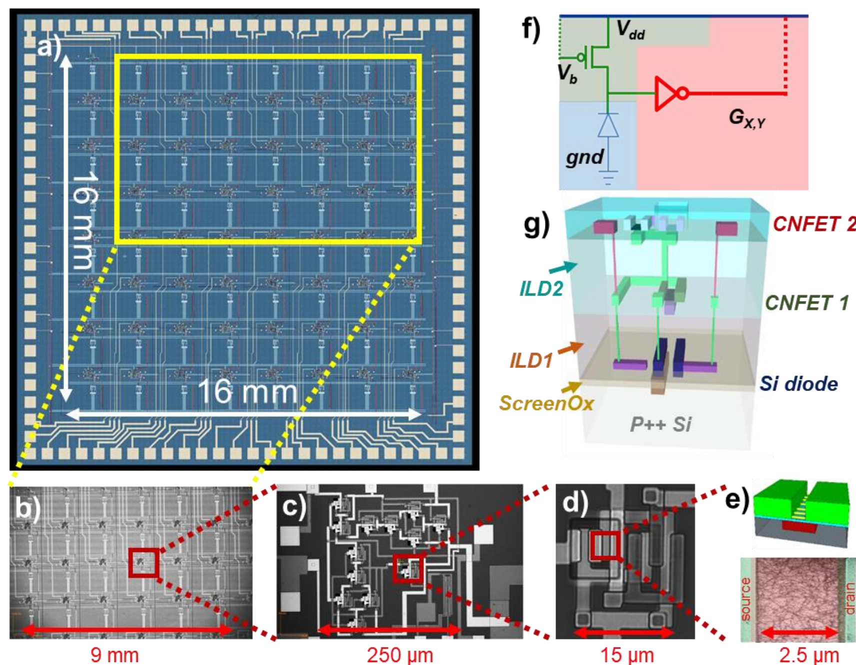
C17-1 Micrograph of 2T-MONOS eFlash macro.

SENSORS, RF, IoT & BIOMEDICAL

3D Imaging System Utilizing Carbon Nanotube Circuitry

MIT will demonstrate a hardware prototype of a monolithic 3D imaging system that integrates computing layers directly in the back-end-of-line (BEOL) of a conventional silicon imager. Such system can transform imager output from raw pixel data to highly processed information. They fabricate 3 vertical circuit layers directly on top of each other: a bottom layer of silicon pixels followed by two layers of CMOS carbon nanotube FETs (CNFETs) that perform in-situ edge detection in real-time, before storing data in memory. This approach promises to enable image classification systems with improved processing latencies.

Paper T2-5 "Monolithic Three-Dimensional Imaging System: Carbon Nanotube Computing Circuitry Integrated Directly Over Silicon Imager,"
T. Srimani et al., Massachusetts Institute of Technology.

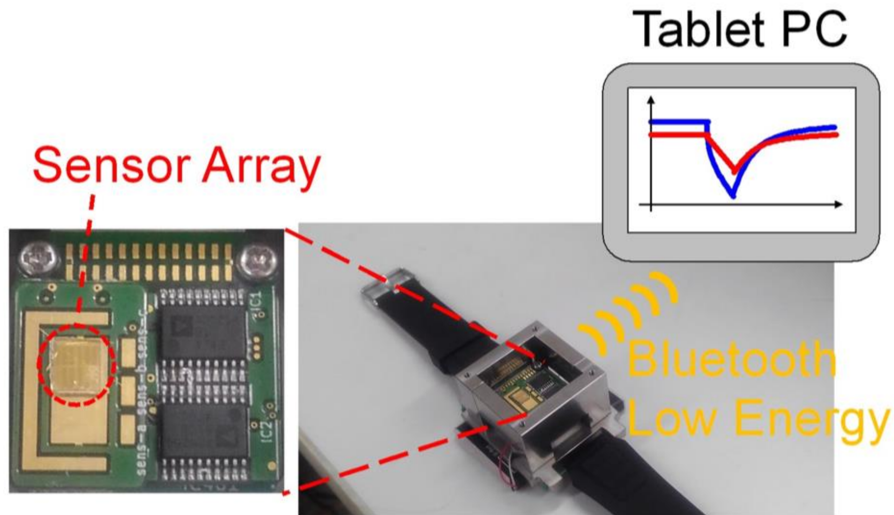


Paper T2-5 3D imager schematics consisting of a Si photodiode (layer 1) and CNFETs (layers 2 and 3).

Low Power Gas Sensor Using Catalytic Metal Nanosheet

Low power gas sensors are useful in IoT applications. Keio Univ. will present integrated sensors for hydrogen and ammonia consisting of catalytic metal nanosheet. The thermal energy necessary for catalytic reactions was provided by on-chip Joule heating instead of the off-chip heaters conventionally used for this application. This is the key to the low power consumption (0.14 mW).

Paper JFS2-3 "Low-Power and ppm-Level Detection of Gas Molecules by Integrated Metal nanosheets,"
T. Tanaka et al., Keio Univ.

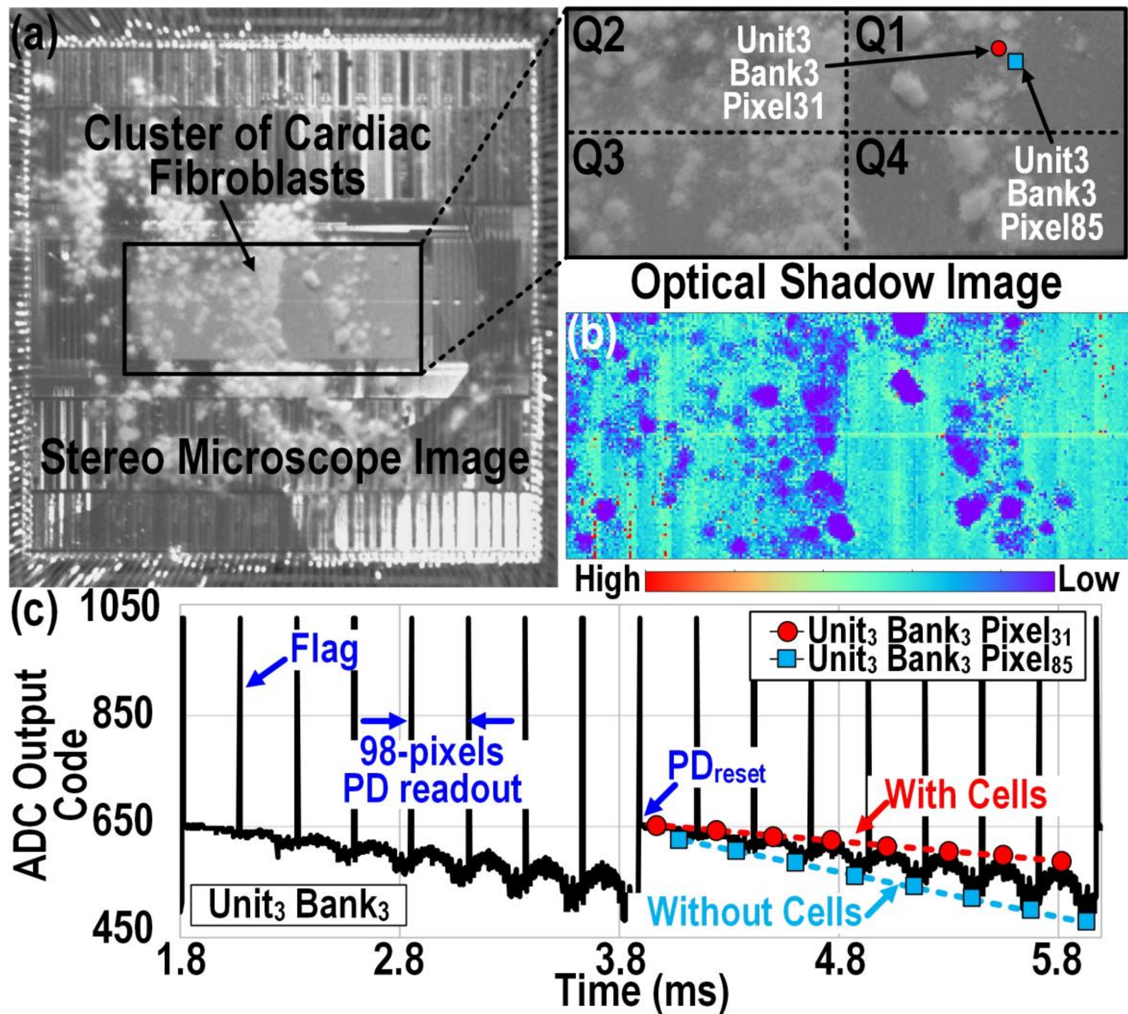


JFS2-3 Gas sensor integrated in a watch type case.

CMOS Multi-modal Cellular Sensor/Stimulator Array

An on-chip cellular image sensor with multi-modal sensing capability is presented by Georgia Institute of Technology. The sensor is capable of capturing optical / potential / 4-point impedance images of cells placed / cultured directly on the sensor. It can also perform electric stimulation onto the cell system. The chip is fabricated using a 130 nm standard CMOS process. The sensor is equipped with 21952 reconfigurable pixels and 1568 pixels can be operated in parallel. Pixel pitch is 16 µm×16 µm, while each pixel has an 8 µm×8 µm gold electrodes and 6 µm×6 µm photo diodes. The electrodes are flexibly addressed to various current injection and impedance measurements. All the functionality of the sensor was characterized. Potential sensing, as well as optical and impedance imaging capabilities were demonstrated in in-vitro experiments with living cells.

Paper C6-3 "A 21952-Pixel Multi-Modal CMOS Cellular Sensor Array with 1568-Pixel Parallel Recording and 4-Point Impedance Sensing,"
D. Jung et al., Georgia Institute of Technology.

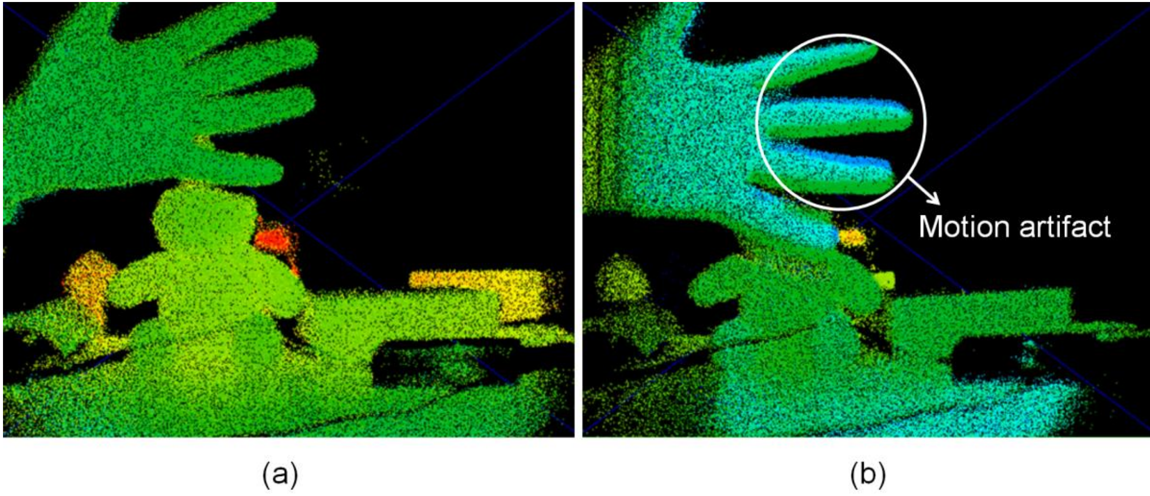


C6-3 (a)Stereo microscope image, (b)measured optical image of on-chip cultured fibroblast and (c)optical time division multiplexing 98 channels ADC output.

Time-of-flight CMOS Image Sensors

Time-of-Flight (ToF) ranging systems are new and promising applications for automotive, human monitoring, and machine vision cameras accelerated by the recent development of high speed and high sensitivity global shutter CMOS image sensors (CIS). Samsung Electronics, Corp. presents a new indirect ToF CMOS image sensor with 640x480 7µm pixels based on a 65nm back-side illumination process. The new sensor achieves motion-artifacts-free operation by simultaneously driving 4-taps and, by a new clock delay control circuit, significantly reduced column fixed pattern phase noise (FPPN) compared with conventional sensors based on 2-taps pixels. The sensor is operable up to 400 cm range with only 0.64 pW/pixel power consumption.

Paper C21-3 "A 640x480 Indirect Time-of-Flight CMOS Image Sensor with 4-tap 7-um Global-Shutter Pixel and Fixed-Pattern Phase Noise Self-Compensation Scheme,"
M.-S. Keel, et al., Samsung Electronics Co., Ltd.



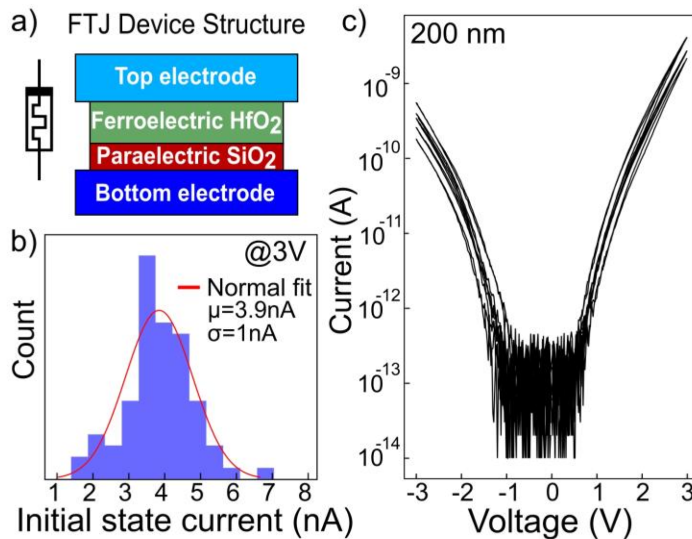
C21-3 Depth map of waving hand, (a)4-tap and (b)2-tap mode.

ARTIFICIAL INTELLIGENCE & QUANTUM COMPUTING

Ferroelectric Tunnel Junction for In-memory Reinforcement Learning System

Toshiba exploits nano-scale ferroelectric tunnel junction (FTJ) memristors with inherent analogue stochastic switching arranged in selector-less crossbars to demonstrate an analogue in-memory reinforcement learning (RL) system, which is capable of learning behavior policies via a hardware-friendly algorithm. The authors will show that commonly undesirable stochastic conductance switching is actually, in moderation, a beneficial property which promotes policy finding via a process akin to random search. They experimentally demonstrate path-finding based on reinforcement, and solve a standard control problem of balancing a pole on a cart via simulation, outperforming similar deterministic RL systems.

Paper T2-4 "In-memory Reinforcement Learning with Moderately-Stochastic Conductance Switching of Ferroelectric Tunnel Junctions."
R. Berdan et al., Toshiba Corp.



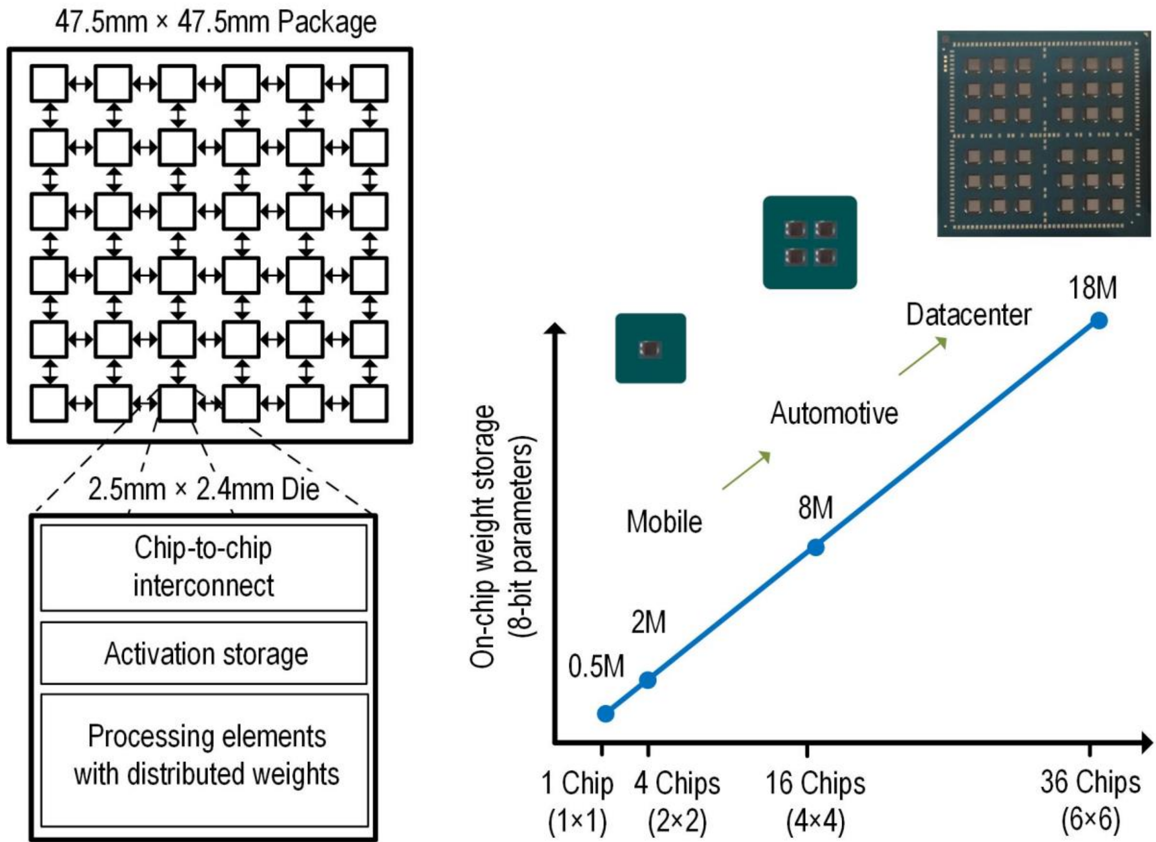
T2-4 FTJ device structure and initial state current distribution.

Scalable, Multi-Chip-Module-based Deep Neural Network Accelerator

Deep neural networks (DNNs) have diverse performance, accuracy, and power targets. Building a dedicated accelerator for each target is often prohibitive due to high design and manufacturing costs. NVIDIA Corp. will present a scalable DNN accelerator consisting of 36 chips connected in a mesh network on a multi-chip-module (MCM) using ground-referenced signaling. While previous accelerators fabricated on a single monolithic die are limited to specific network sizes, the proposed architecture enables flexible scaling for efficient inference on a wide range of DNNs, from mobile to data center domains. The 16nm prototype achieves 1.29 TOPS/mm2, 0.11pJ/op, 4.01TOPS peak performance for a 1-chip system, and 127.8 peak TOPS and 2615 images/s ResNet-50 inference for a 36-chip system.

Paper C24-1 "A 0.11 pJ/Op, 0.32-128 TOPS, Scalable, Multi-Chip-Module-based Deep Neural Network Accelerator with Ground-Reference Signaling in 16nm,"

B. Zimmer, et al., NVIDIA Corp.



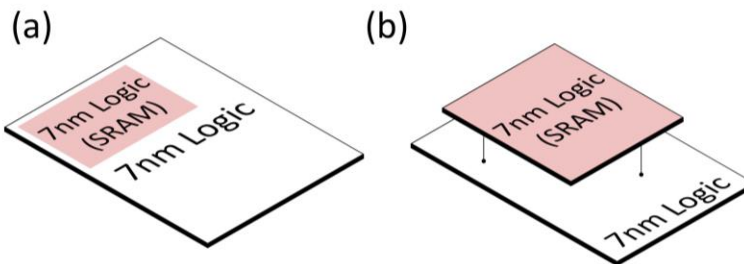
C24-1 System overview of the proposed MCM-based DNN accelerator.

2.5D/3D Integrations

System on Integrated Chips (SoIC^{TM})

TSMC will reveal System on Integrated Chips (SoICTM), an innovative 3D heterogeneous integration technology manufactured in front-end of line with known-good-die. Chiplets integration of devices with SoIC^{TM} illustrates its advantages in high bandwidth density and high power efficiency, as compared with 2.5D and conventional 3D-IC with micro-bump/TSV.

Paper T2-3 "3D Multi-chip Integration with System on Integrated Chips (SoICTM)," C. C. Hu et al., TSMC.

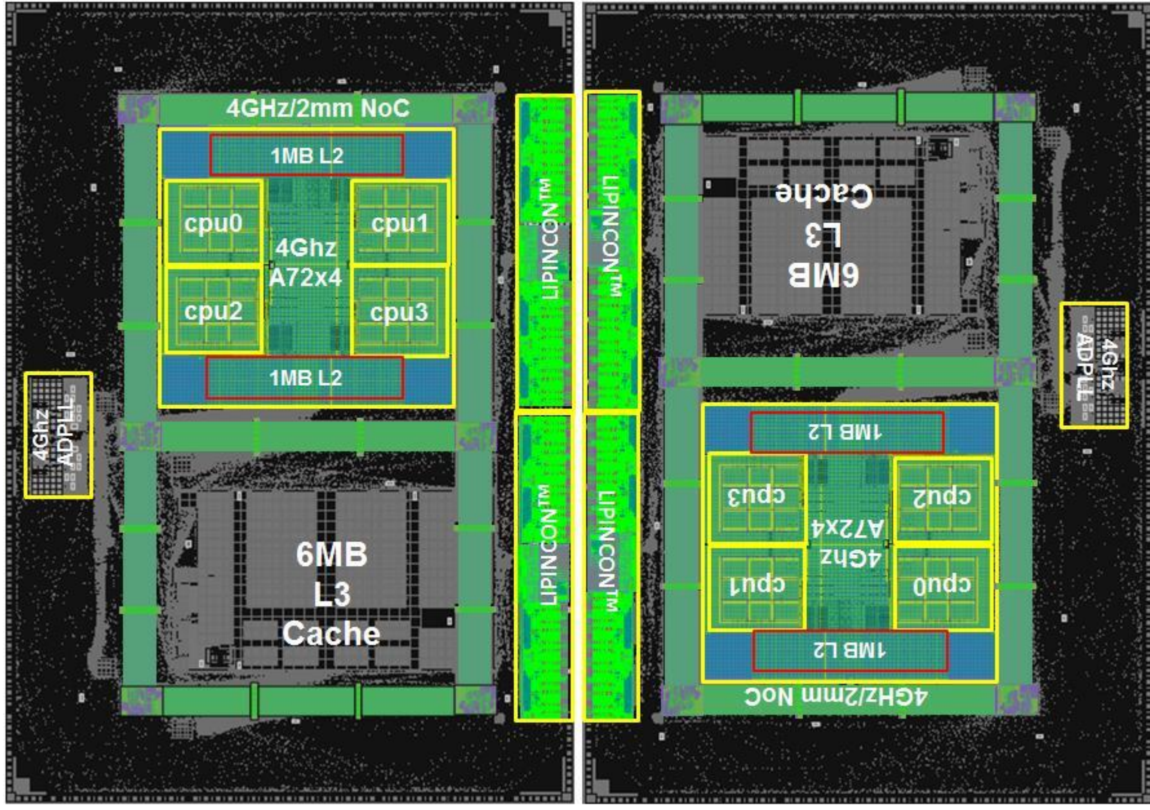


T2-3 SoIC interconnect is compared with typical 2.5D or 3D-IC.

Dual-Chiplet High-Performance Computing Processor in 7nm Process

A chiplet-based design becomes increasingly important to realize large-scale and high-performance processors with better yield and lower product cost. This year, TSMC will present a dual-chiplet high-performance computing (HPC) processor implemented in 7nm process with CoWoS® technology. Each chiplet has 4 ARM Cortex®-A72 cores operating at 4GHz at turbo voltage condition, and the on-die inter-core mesh interconnect operates above 4GHz. The inter-chiplet connection interface, called Low-Voltage-In-Package-INterCONnect (LIPINCON™), provides 0.56pJ/bit power efficiency, 1.6Tb/s/mm2 bandwidth density, and 320GB/s bandwidth. This work demonstrates that the chiplet-based large-scale and high-performance processor is applicable for modern high-performance computing era.

Paper C3-1 "A 7nm 4GHz Arm-core-based CoWoS Chiplet Design for High Performance Computing," M.-S. Lin, et al., TSMC.



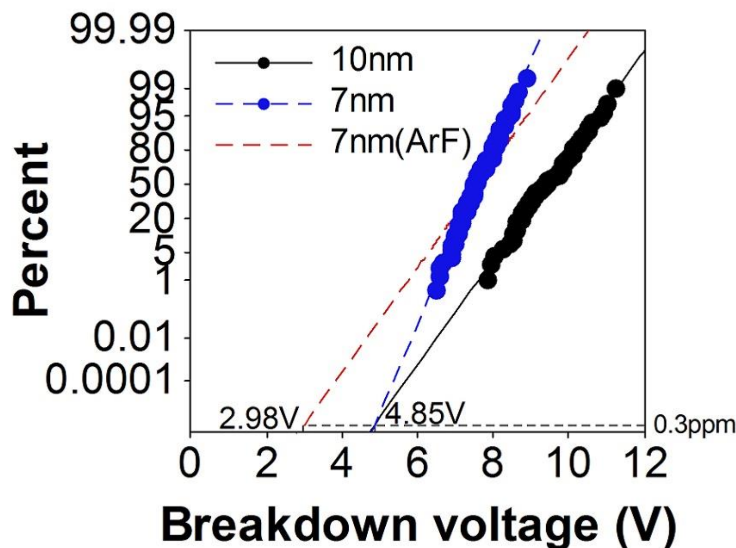
C3-1 Dual chiplet Floorplan.

ADVANCED PLATFORM & SOCs

Enhanced Reliability of 7nm Technology by EUV

Samsung Electronics will report the reliability characterization of 7nm FinFET technology fabricated with EUV. The single EUV patterning of MOL and BEOL resulted in significantly improved reliability distribution when compared to the previous nodes with multiple patterning techniques. They successfully demonstrated product reliability including SRAM, Logic HTOL, and SER. These results indicate that 7nm technology with EUV is ready for high volume manufacturing.

Paper T2-1 "Enhanced Reliability of 7nm Process Technology featuring EUV,"
K. Choi et al., Samsung Electronics.

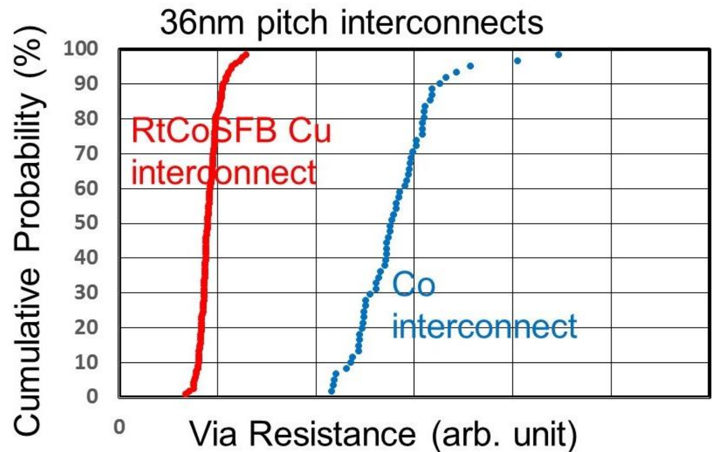


T2-1 The VBD slope of gate-to contact Vramp distribution showing significant improvement with EUV process.

Cu Interconnects with RtCoSFB for 7nm and Beyond

IBM will demonstrate EM and TDDB reliability of Cu interconnects with a barrier/wetting layer as thin as 2 nm employing a PVD-reflowed through-Co self-forming barrier (RtCoSFB). The resulting Cu EM lifetime with RtCoSFB is 2000X longer than Cu interconnects with a standard scaled barrier/wetting layer, and is equivalent to pure Co interconnects. Although Cu interconnects with RtCoSFB enables lower line resistance and via resistance than Co interconnects, the annealing process for RtCoSFB causes Cu agglomeration at dual damascene line-end vias, leading to poor via-chain yield. They identify resolving this geometry-sensitive via-fill problem as key to extending Cu manufacturability to 7 nm and beyond.

Paper T2-2 "Technology challenges and enablers to extend Cu metallization to beyond 7 nm node,"
T. Nogami et al., IBM Research.

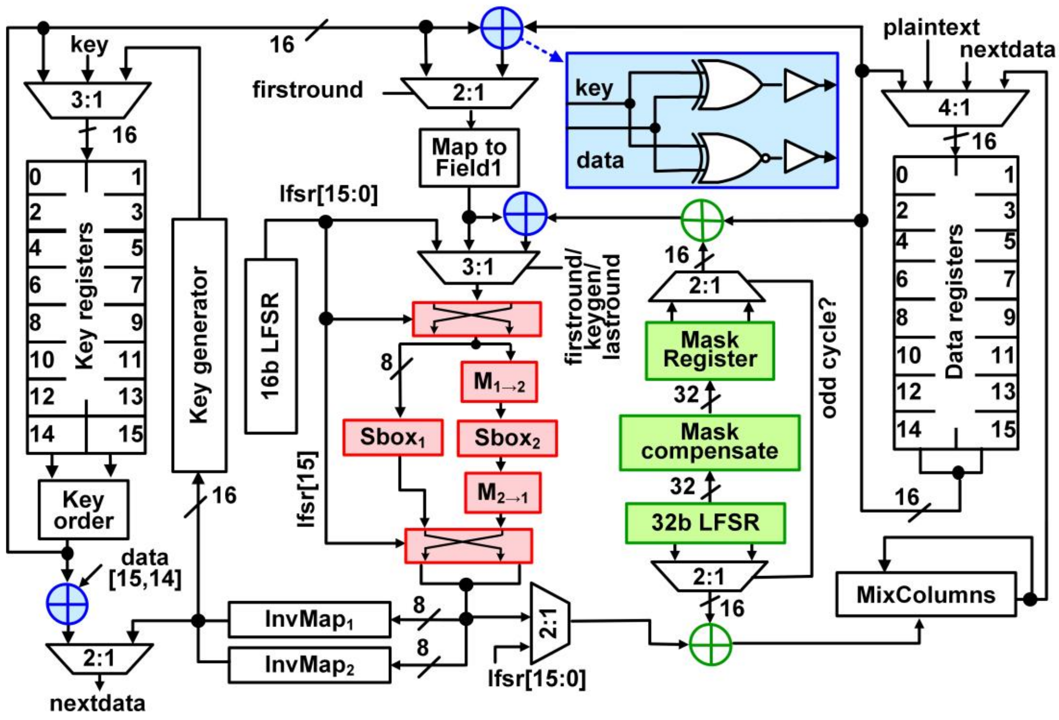


T2-2 Cu with RtCoSFB enables 1/3x via-R of that of Co interconnects.

Light-Weight Secure AES Accelerator in 14nm CMOS

Hardware security is one of the most critical issues in IoT where a malicious attacker exploits physical security holes of cryptographic accelerators to steal valuable information processed in SoCs. A side-channel attack is one well-known physical attack based on a statistical analysis on power consumption of the accelerators. This year, Intel will be presenting a light-weight secure AES accelerator fabricated in 14nm CMOS. This accelerator exhibits 1200x stronger resiliency against the power-analysis side-channel attack. The design features random heterogeneous Sboxes shuffling, linear masked MixColumns, and dual-rail key addition to significantly suppress the correlation between the power consumption and a secret key without significant hardware overhead. Measured results demonstrate attack resistant operation over 12M-times encryption traces. The hardware overhead is only 23%, 28%, and 0.7% in power, area, and performance, respectively.

Paper C20-1 "A 4900um2 839Mbps Side-Channel Attack Resistant AES-128 in 14nm CMOS with Heterogeneous Sboxes, Linear Masked MixColumns and Dual-Rail Key Addition,"
R. Kumar, et al., Intel



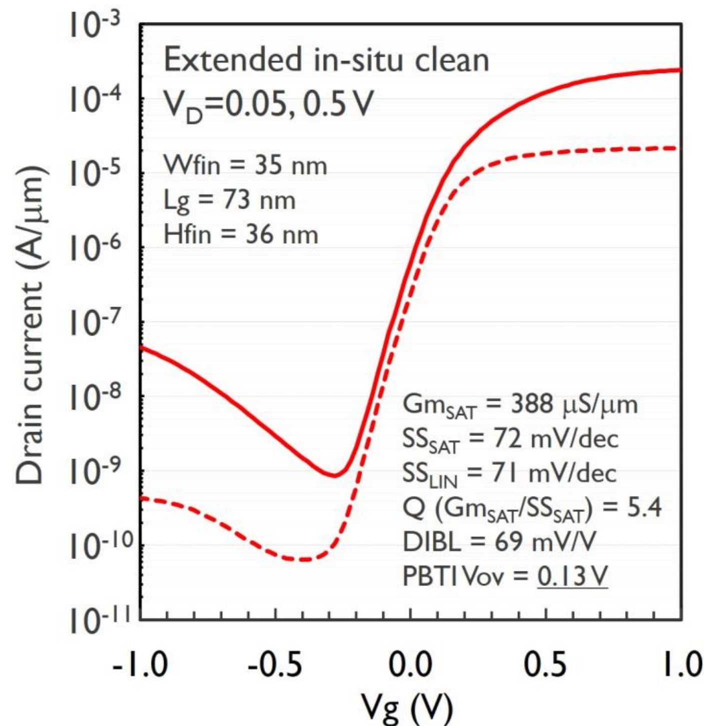
C20-1 Side-channel attack resistant AES-128 organization.

ADVANCED TRANSISTOR TECHNOLOGY

Record High Device Performance and Reliability in Ge nFinFET

Imec demonstrates that the SiO_2 dummy gate oxide (DGO) deposition and removal processes are two important knobs to improve Ge nFinFETs characteristics. By suppressing the Ge channel surface oxidation during DGO deposition, increasing mobility with decreasing fin width is obtained, whereas PBTI reliability, D_{it} of scaled fin as well as high-field mobility are improved by extending the DGO removal process, resulting in the record high Gm/SS of 5.4 at 73 nm L_g.

Paper T9-1 "A record Gmsat/SSsat and PBTI reliability in Si-passivated Ge nFinFETs by improved gate stack surface preparation,"
H. Arimura ,et al., imec.

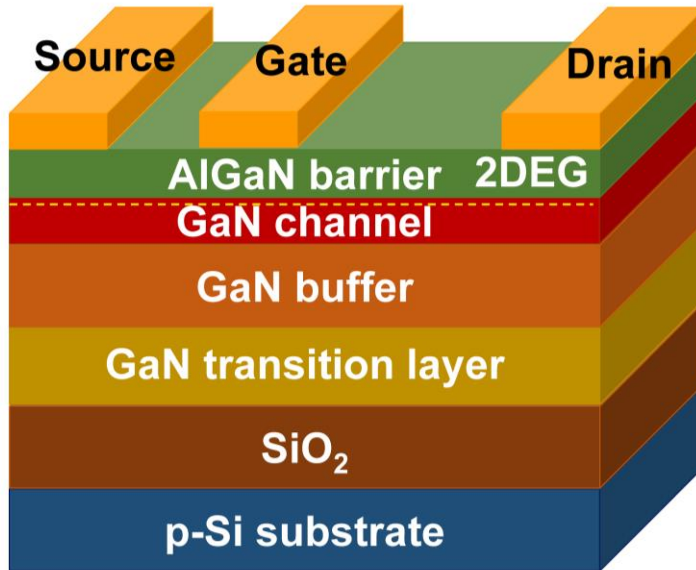


T9-1 ID-VG curves obtained from 35 nm-wide Ge nFinFET.

High Breakdown Voltage GaN HEMT Realized by GaN-on-Insulator(GNOI)-on-Si Wafer

GaN-on-Si technology is one of the promising candidates for next-generation power electronics applications. Singapore-MIT Alliance for Research and Technology (SMART) will present 200-mm GaN-on-Insulator (GNOI)-on-Si wafer by using an epitaxial layer transfer technique. Fabricated HEMTs exhibit a record high off-state breakdown voltage up to 2200 V and high figure-of-merit of BVoff2/Ron,sp up to 1.87 GW/cm2.

Paper T19-1 "GaN HEMTs with Breakdown Voltage of 2200 V Realized on a 200 mm GaN-on-Insulator(GNOI)-on-Si Wafer,"
Z. Liu, et al., Singapore-MIT Alliance for Research and Technology (SMART).

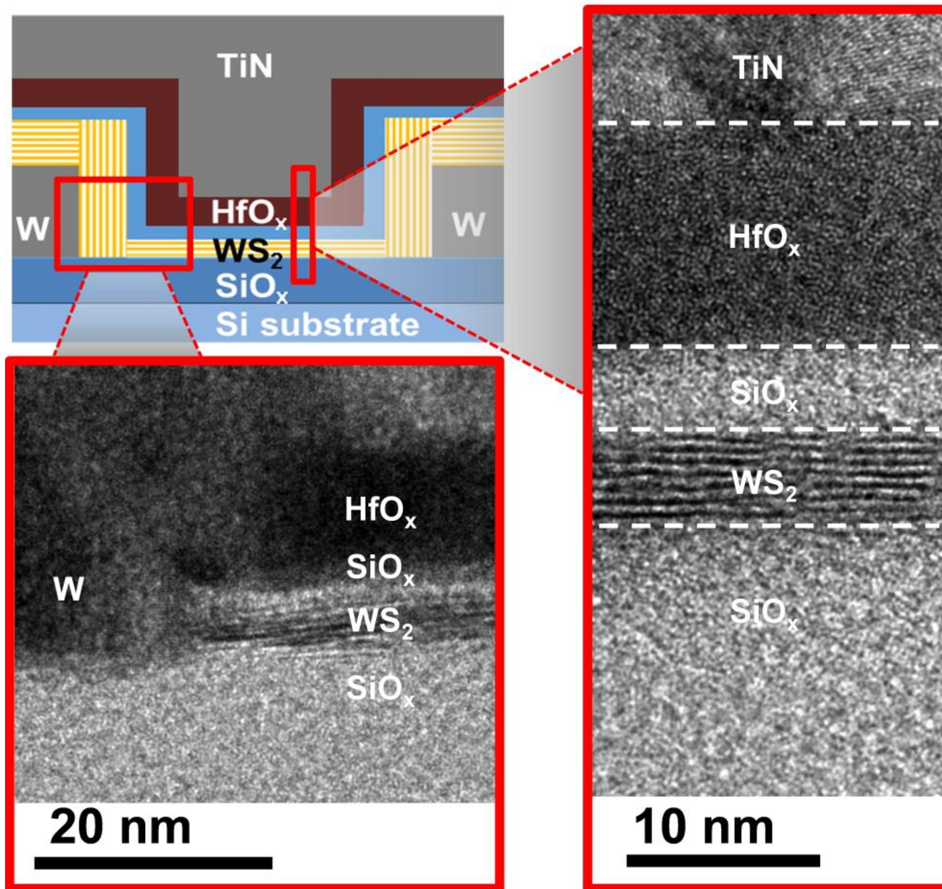


T19-1 Proposed GaN HEMT structure.

WS_2 Channel pFET Demonstration Using Channel Area-Selective CVD Growth Directly on SiOx/Si Substrate

Selective-area growth of 2D materials is a key technology for volume manufacturing of 2D-channel transistors with suppressed short channel effects. TSMC, Taiwan Semiconductor Research Institute will present the first top-gate WS_2 p-channel FETs fabricated on SiOx/Si substrate using selective-area CVD growth. Fabricated 40-nm-gate-length WS_2 FETs show on/off ratio of 10^6, a subthreshold slope of 97 mV/dec, and nearly zero DIBL.

Paper T19-2 "First demonstration of 40-nm channel length top-gate WS2 pFET using channel area-selective CVD growth directly on SiOx/Si substrate,"
C.-C. Cheng, et al., TSMC.

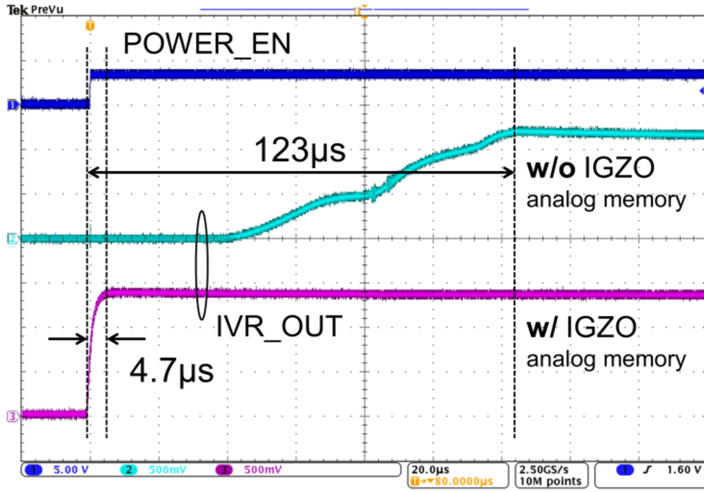


T19-2 Cross sectional image of proposed WS2 pFET.

IGZO-based MCU for Normally-Off Computing

One of the innovative low power technologies introduced this year is an IGZO-based MCU designed for normally-off computing. The MCU is implemented using 60-nm IGZO process, retains data during power gating in both of its processing unit and memory, realizing only 1 clock(21ns) cycle backup time and a speedy 4.49us restoration. 880nW ultra low standby power is realized by the low leakage current of the IGZO. This work proves that the IGZO technology can be one of the candidates to accelerate widespread use of IoT and edge AI applications that need low power and fast wakeup.

Paper C5-1 "A 48 MHz 880-nW Standby Power Normally-Off MCU with 1 Clock Full Backup and 4.69-us Wakeup Featuring 60-nm Crystalline In-Ga-Zn Oxide,"
T. Ishizu, et al., Semiconductor Energy Laboratory Co., Ltd.



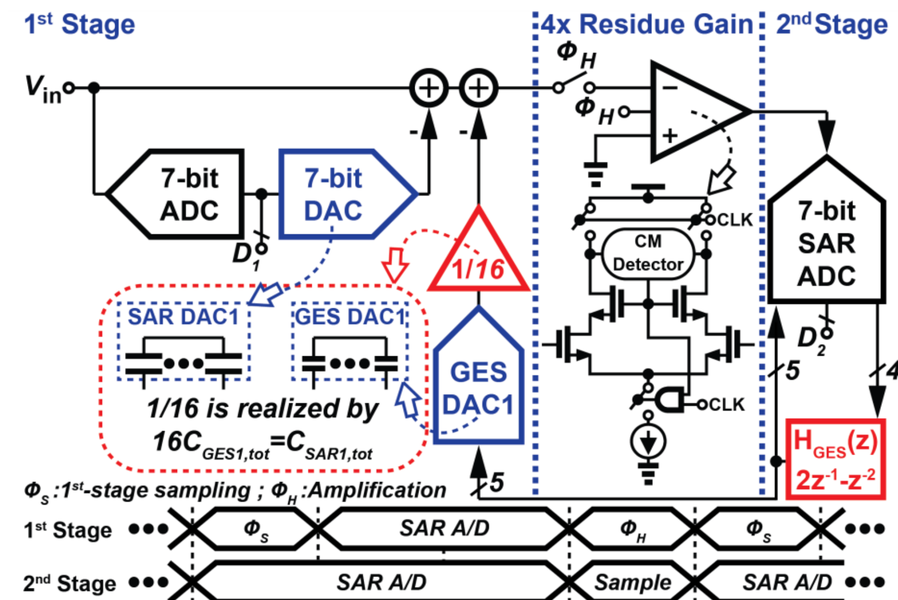
C5-1 Settling time of integrated voltage regulator w/ or wo/ IGZO analog memory.

WIRELINE & DATA CONVERTER CIRCUITS

75.8dB-SNDR Pipeline SAR ADC with 2nd-order Interstage Gain Error Shaping

Pipeline SAR ADCs widely used for high-speed applications are sensitive to interstage gain error. A paper from The University of Texas at Austin presents a low-cost gain error shaping (GES) technique that suppresses the in-band interstage gain error in pipeline SAR ADCs. The technique works for both closed-loop and open-loop interstage amplification. The pipeline SAR ADC with proposed 2nd-order GES technique fabricated in 40nm CMOS achieves 75.8dB SNDR over 12.5MHz BW while operating at 100MS/s and consuming 1.54mW. The GES-related hardware occupies less than 2% of the core area.

Paper C7-1 "A 75.8dB-SNDR Pipeline SAR ADC with 2nd-order Interstage Gain Error Shaping," C-K. Hsu, et al., The University of Texas at Austin

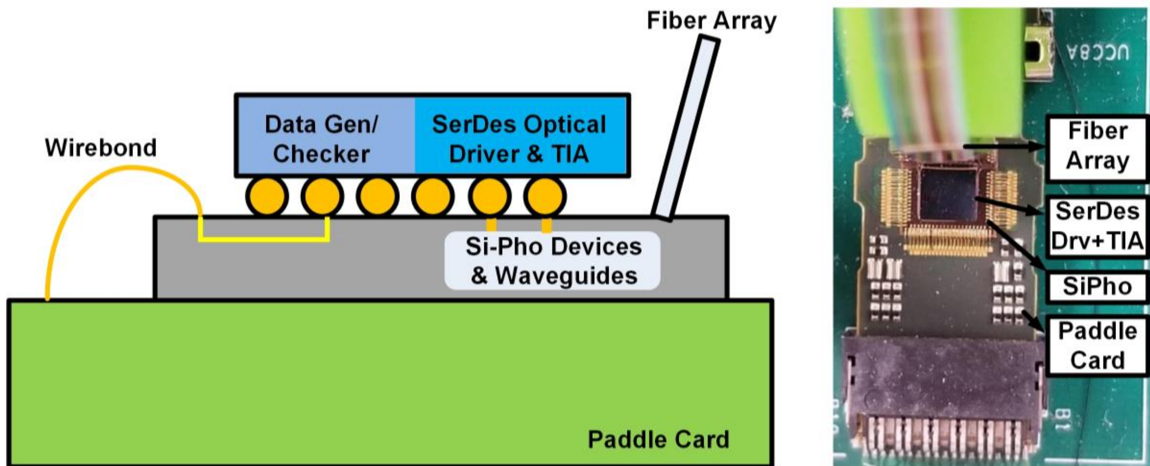


C7-1 Block and timing diagram of the ADC with 2nd-order GES.

50Gb/s Hybrid Integrated Si-Photonic Optical Link

Silicon Photonics is a key technology to achieve high bandwidth-density and energy-efficient above electrical signaling. This year, Xilinx will present a 50Gb/s hybrid integrated Si-Photonic optical link. The Tx and Rx fabricated in 16nm CMOS use T-coils to improve modulation efficiency of the electro-absorption modulator and to reduce the noise of the transimpedance amplifier. The link sensitivity is -10.9dBm optical modulation amplitude at BER<10^{-12} and it consumes 4.31pJ/bit at 50Gb/s with the laser.

Paper C16-1 "A 50Gb/s Hybrid Integrated Si-Photonic Optical Link in 16nm FinFET," M. Raj, et al., Xilinx, Inc.



C16-1 Optical transceiver (driver and TIA) over Si-photonics: Block diagram (left) and photo(right).

112Gb/s PAM4 Receiver for Long-Reach Channels in 10nm Process

The world's first integrated 112Gb/s PAM4 receiver was presented in last Symposium on VLSI Circuits. This year Intel enhanced the stage-of-the-art by demonstrating longer-reach and small area 112Gb/s PAM4 receiver. In addition to 64-way interleaved SAR-ADC and 2step cascade sampling topology in the decision block, it utilizes LC passive element based Q-shaping equalizer technique in the analog front-end, enabling low noise and strong equalization performance. The work is implemented in 10nm process, and evaluated using 112Gb/s PAM4 PRBS-31 input data pattern. The receiver including 16tap FFE and 1tap DFE in the on-chip DSP achieved a pre-FEC raw BER of $< 10^{-6}$ over a channel with 35dB loss, while consuming 0.281mm^2 active area.

Paper C22-1 "112Gb/s PAM4 ADC Based SERDES Receiver for Long-Reach Channels in 10nm Process,"
Y. Krupnik, et al., Intel

C22-1 Block diagram of the proposed receiver.

Low Power Continuous-Time Zoom ADC for Audio Applications

The continuous-time (CT) A/D converter using the zoom architecture will be presented by Delft University of Technology. In the field of mobile audio applications, a data conversion system must achieve low power and high resolution. Conventional high-precision A/D converters put a heavy burden on the input and reference drivers because of the capacitive load and sampling. The proposed continuous-time zoom A/D converter has resistive input impedance and therefore can relax such requirement for peripheral drivers. This A/D converter achieves high accuracy conversion such as 108.5 dB dynamic range and 106.4 dB SNDR, while consuming 618 μW.

Paper C19-1 "A Low Power Continuous-Time Zoom ADC for Audio Applications," B. Gönen, et al., Delft University of Technology.

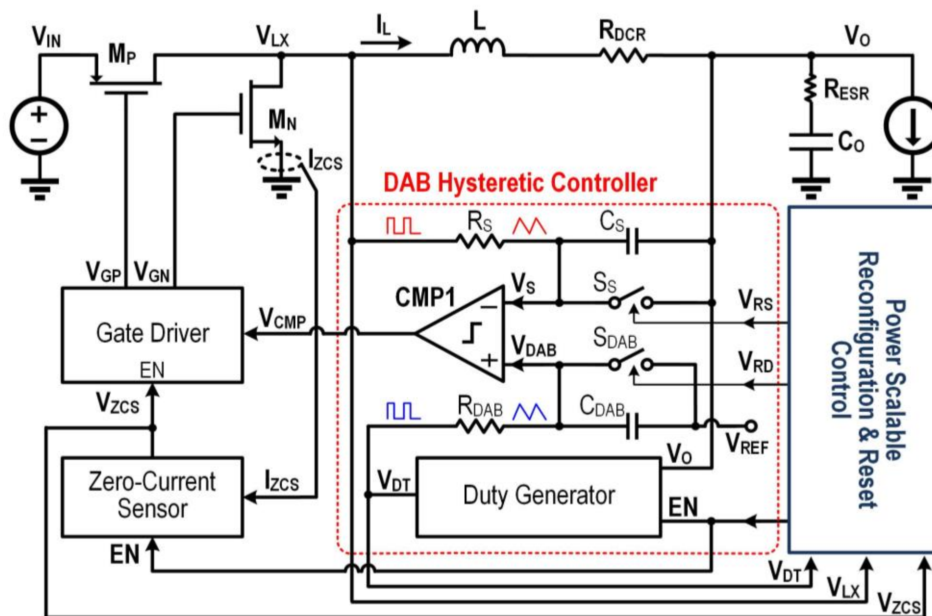
C19-1 Simplified schematic of the proposed CT zoom ADC.

POWER MANAGEMENT CIRCUITS

DC-DC Converter for 5G IoTs

Power converters in 5G IoTs are expected to handle abrupt power flips between zero and maximum power with ultra-fast transient response, deliver a wide range of power efficiently, and retain low system profile and high power density. The Univ. of Texas at Dallas will present a double adaptive bound (DAB) hysteretic control power converter designed for 5G IoTs. In response to 1A/3ns load step-up/step-down, it achieves 1% t_{settle} of 247ns/387ns, thanks to the DAB control. This is 6X faster than the best published work on 0.18μm CMOS. It also achieves >80% efficiency over 99.9% of 2.5W full power range. Highly efficient design leads to the highest reported chip power density of 14.3W/mm^2.

Paper C15-3 "A 10-MHz 14.3W/mm2 DAB Hysteretic Control Power Converter Achieving 2.5W/247ns Full Load Power Flipping and above 80% Efficiency in 99.9% Power Range for 5G IoTs,"
K. Wei, et al., The University of Texas at Dallas



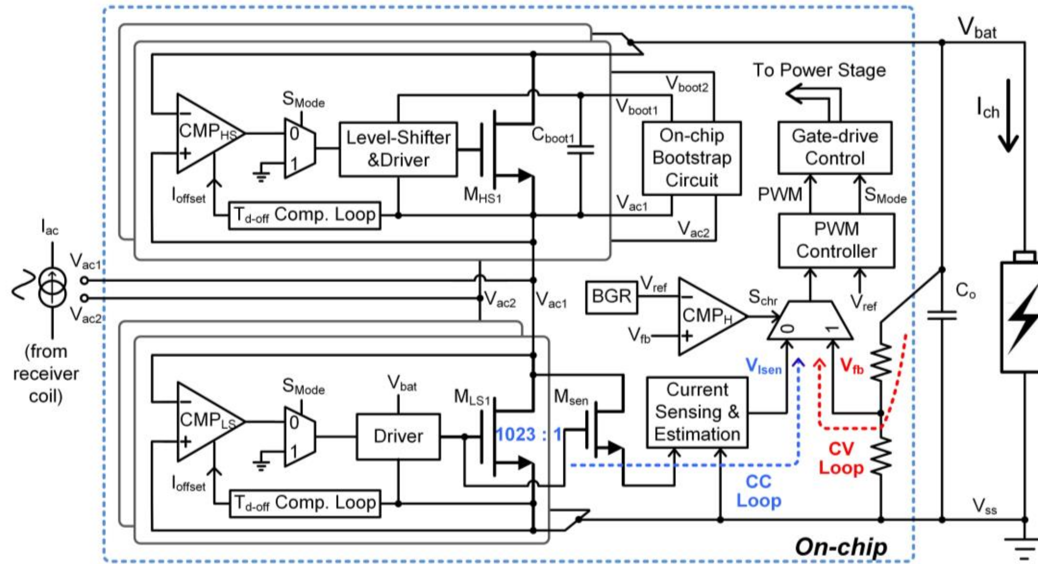
C15-3 Block diagram of the proposed DC-DC converter.

Highly Efficient and Compact Single-Stage Wireless Charger

CC-CV (constant-current and constant-voltage) battery charging is already widely used, but implementing area and power efficient wireless CC-CV for 1A-class applications has been a challenge. USTC and HKUST succeeded to perform three functions - voltage rectification, regulation and CC-CV charging - in their compact single-stage power converter along with an integrated bootstrap capacitor on an 8mm^2 IC in 0.35um CMOS. The measured peak efficiency is 92.3% and 91.4% when the charging currents are 1A and 1.5A, respectively.

Paper C26-1 "A 6.78MHz 92.3%-Peak-Efficiency Single-Stage Wireless Charger with CC-CV Charging and On-Chip Bootstrapping Techniques,"

L. Cheng, et al., University of Science and Technology of China.



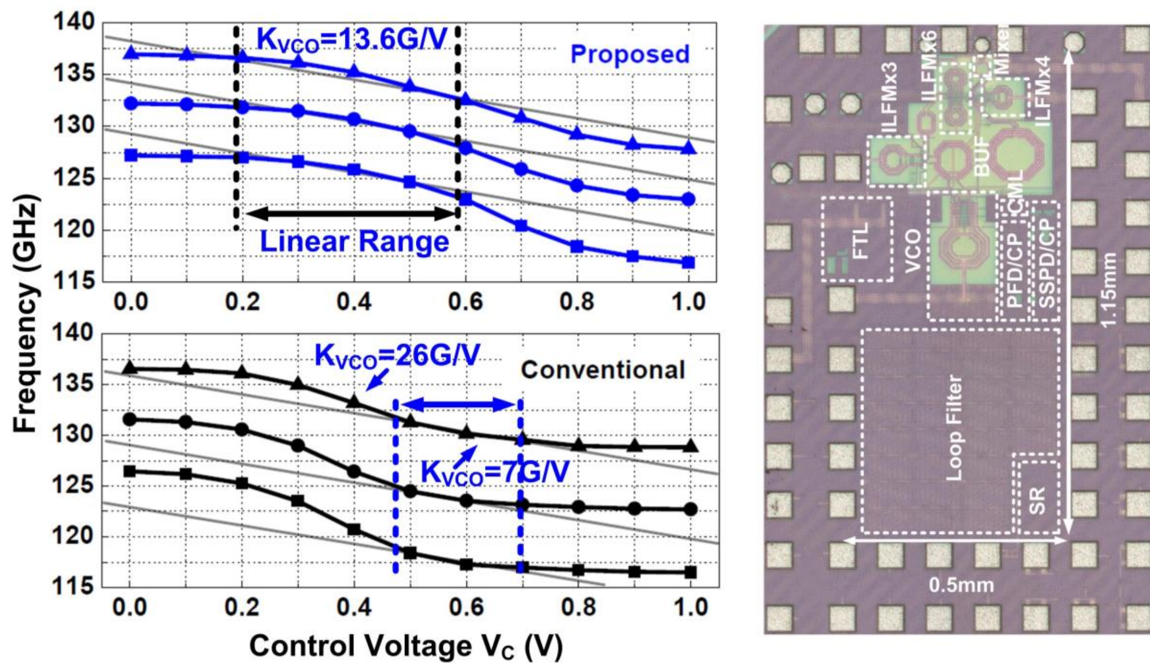
C26-1 Block diagram of the proposed charger.

FREQUENCY GENERATION

270-GHz Fully-Integrated Frequency Synthesizer in 65nm CMOS

A wideband, low-phase-noise, and accurate sub-THz source is essential in a coherent rotational spectroscopy system as well as terahertz wireless communication. HKUST will report a sub-THz frequency synthesizer implemented in cost-effective 65nm CMOS, which can generate 61.2-to-100.8GHz, 122.4-to-136.8GHz, and 198.5-to-273.6GHz by using a cascaded ILFM (Injection-Locked Frequency Multiplier) chain. An output power of -11dBm and DC-to-RF efficiency of 0.16% at a carrier of 211GHz is achieved.

Paper C4-2 "A 270-GHz Fully-Integrated Frequency Synthesizer in 65nm CMOS,"
X. Liu, et al., Hong Kong University of Science and Technology



C4-2 Measured frequency tuning range of an injection locked frequency multiplier and chip micrograph of the proposed frequency synthesizer.