Technical Highlights from the 2023 Symposium on VLSI Technology and Circuits

The 2023 Symposium on VLSI Technology and Circuits is a premiere international conference that records the pace, progress, and evolution of micro/nano integrated electronics, scheduled for June 11-16, 2023. The Symposium will be held in-person at the Rihga Royal Hotel, Kyoto Japan to foster networking opportunities.

The Symposium's overall theme, **"Rebooting Technology and Circuits for a Sustainable Future"** integrates advanced technology developments, innovative circuit designs, and the applications that they enable as part of our global society's transition to a new era of smart connected devices, infrastructure and systems that change the way humans interact with each other.

The following are some of the highlight papers that address this theme:

Joint Technology and Circuits Highlights

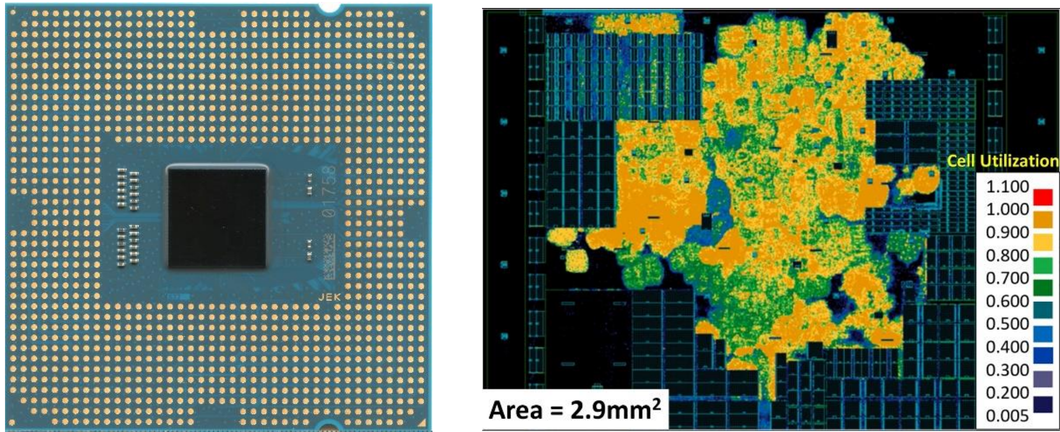
Here are some joint highlight papers pushing a combined step forward in technology and circuits:

Processors

"E-Core Implementation in Intel 4 with PowerVia (Backside Power) Technology" – Intel Corp. (Paper T1-1)

Intel reports a high-yielding backside power delivery technology, PowerVia Technology*, and Intel E-Core Implementation in PowerVia Technology. PowerVia Technology is a novel innovation to extend Moore's Law scaling and enables standard cell utilization exceeding 90% in large areas of the core while showing more than 5% frequency benefit in silicon due to reduced IR drop. Successful post-silicon debug is demonstrated with slightly higher but acceptable throughput times. The thermal characteristic of the PowerVia test chip is in line with higher power densities expected from logic scaling.

* PowerVia Technology is presented in Session T6-1.



Figures: (Left) Die photo of Intel 4 with PowerVia. (Right) Cell density plot.

Devices and Accelerators for Machine Learning

"Chip Demonstration of a High-Density (43Gb) and High-Search-Bandwidth (300Gb/s) 3D NAND Based In-Memory Search Accelerator for Ternary Content Addressable Memory (TCAM) and Proximity Search of Hamming Distance" – Macronix International Co., Ltd. (Paper T15-1)

Macronix demonstrates a high-density (43Gb) in-memory search chip based on a 96-layer 3D NAND product. The novel String-Select Line design achieved 300Gb/s search bandwidth with a measured power of less than 400mW, accelerating both exact TCAM and proximity Hamming-distance searches.

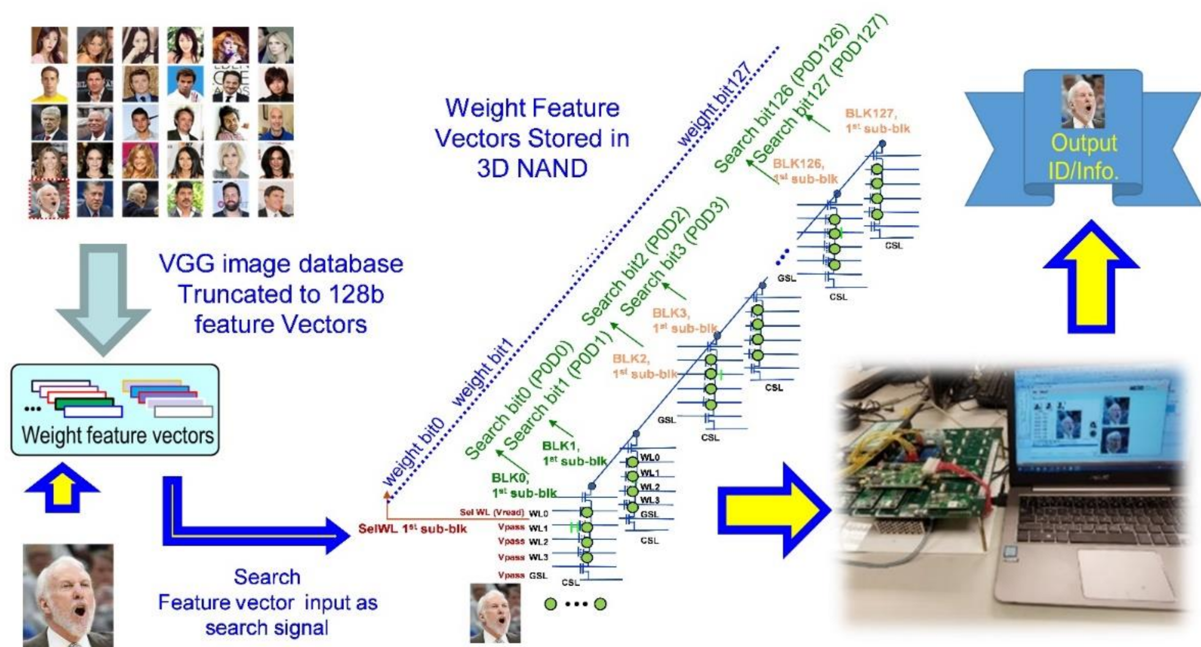


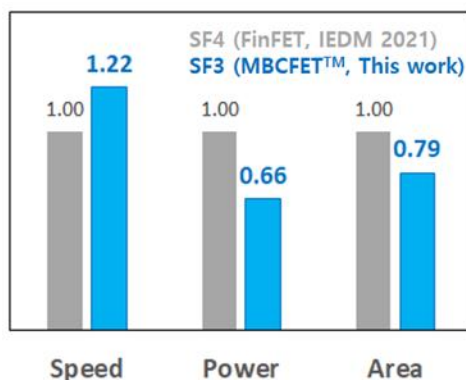
Figure: The processing flow of IMS operations. They use public face recognition (VGGGace2 database) for the demonstration. The feature vectors (128b in this case) are extracted and stored in the chip. For data querying, the search data is input to the chip, and the chip can directly compute which BL address matches the result.

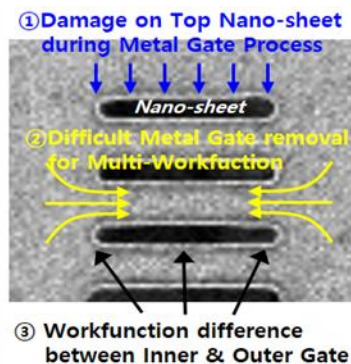
Technology Highlights

Advanced CMOS Technology

“World's First GAA 3nm Foundry Platform Technology(SF3) with Novel Multi-Bridge-Channel-FET (MBCFET) Process” – Samsung Electronics Co. Ltd. (Paper T1-2)

This paper reports technical details on the highly anticipated transition from FinFET to Gate-All-Around (GAA) transistor architecture. Samsung reveals the world’s first GAA 3nm Foundry Platform Technology, SF3, which provides 22% speed and 34% power improvements and 0.79x logic area reduction over the 4nm FinFET platform. SF3 technology is the upgraded version of the industry first mass-produced GAA (SF3E). Through Multi-Bridge-Channel FET (MBCFET™) unique process design, SF3 provides various nanosheet widths with comparable performance at fixed standard cell height to significantly enhance chip-level power-performance matrix over the FinFET platform.



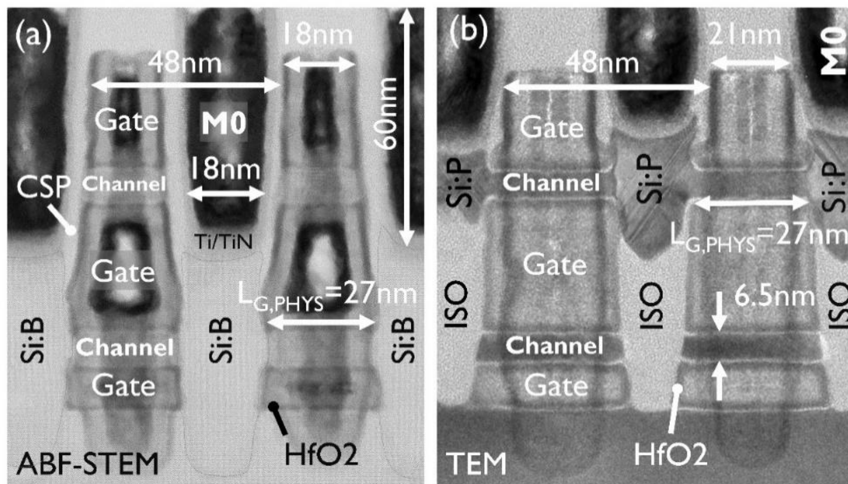


Figures: (Left) Performance/Power/Area (PPA) comparison between the previous 4nm FinFET (SF4) and the world first 3nm SF3 GAA (MBCFET™). (Right) Cross section of MBCFET™ highlighting unique process optimization.

Advanced CMOS Technology

“Nanosheet-based Complementary Field-Effect Transistors (CFETs) at 48nm Gate Pitch and Middle Dielectric Isolation to enable CFET Inner Spacer Formation and Multi-Vt Patterning” – imec (Paper T1-3)

Researchers at imec report on silicon nanosheet monolithic Complementary Field-Effect Transistors (CFETs) at an industry-relevant gate pitch of 48nm with source/drain (SDs) and SD contacts formed for either bottom or top devices. SD epi patterning at 30nm vertical N-P space and high-aspect-ratio SD contact formation are successfully demonstrated. The monolithic CFETs have a good subthreshold swing of 70mV/dec. for the NFETs and 75mV/dec. for the PFETs. Middle dielectric isolation (MDI) formed by SiGe replacement processing is introduced as an enabler for monolithic CFET inner spacer formation and multi-V_T patterning.

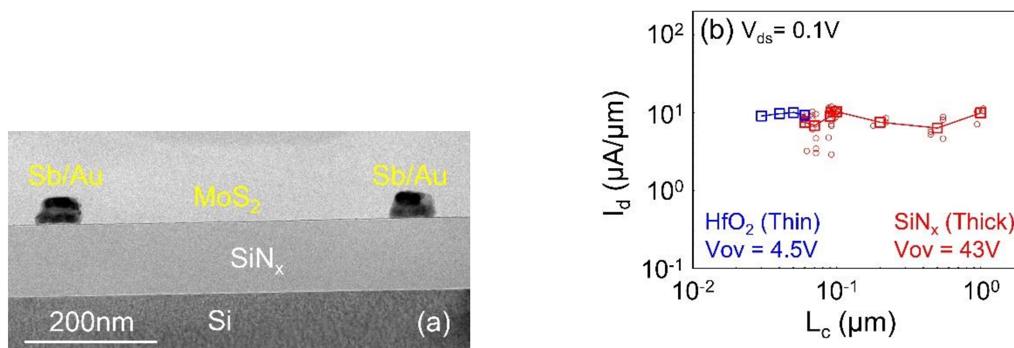


Figures: Cross sectional images for (a) bottom pFET and (b) top nFET.

Advanced Process and Materials

"Scaled Contact Length with Low Contact Resistance in Monolayer 2D Channel Transistors" –Taiwan Semiconductor Manufacturing Company, National Yang Ming Chiao Tung University, Nanjing University, National Cheng Kung University (Paper T1-4)

A research collaboration led by TSMC demonstrates a scenario for low contact resistance at scaled contact lengths in the Sb-MoS_2 system. This is a critical enabler for realizing extremely scaled logic transistors using two-dimensional transition metal dichalcogenides (2D TMDs). 2D TMDs are expected to enable extremely scaled logic transistors and such aggressive scaling also requires scaling of the contact length. The proposed TCAD model was verified with experimental data, and the good narrow-gap R_c of 250 Ω-µm can be extrapolated at contact lengths down to 15 nm.



Figures: Transmission Electron Microscopy (TEM) images of contacts (a) and on-current density over the wide range of L_c.

Advanced Process and Materials

"Contact Cavity Shaping and Selective SiGe:B Low-Temperature Epitaxy Process Solution for sub10^{-9} ohms.cm^2 Contact Resistivity in Nonplanar FETs" – Applied Materials, IBM Semiconductor Technology Research (Paper T1-5)

Applied Materials and IBM collaborate to develop a contact cavity shaping process allowing an active boron doping level of $2x10^{21}$ atoms/cm^3. They co-optimize a reactive ion etch (RIE) and selective highly doped SiGe:B epitaxial process in the contact module on 300 mm wafers. They demonstrate a record low transistor contact resistance of 11ohm.µm. They realize a device effective on-current performance gain of 44% for medium and 19% for leading edge transistors.

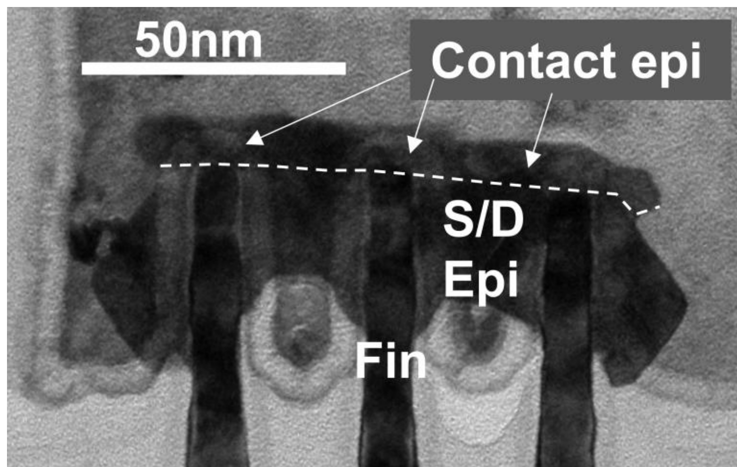


Figure: TEM cross section of FinFET device along the contact trench, showing a continuous growth of the contact epitaxial layer.

Advanced Process and Materials

Beyond 10um Depth Ultra-High Speed Etch Process with 84% Lower Carbon Footprint for Memory Channel Hole of 3D NAND Flash over 400 Layers. – Tokyo Electron Miyagi Ltd. (Paper T3-2)

Authors at Tokyo Electron Miyagi develop a novel etching process for high aspect hole patterning using cryogenic temperature and novel carbon-less chemistry for 3D NAND flash memory devices. This technology shows 10µm etch depth capability and quite short process time (33 minutes) with 84% carbon footprint reduction of greenhouse gases. Excellent etch profile is achieved.

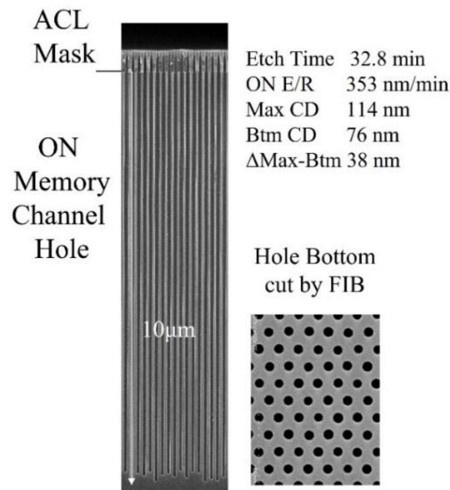
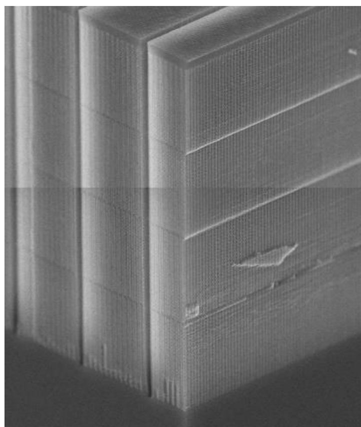


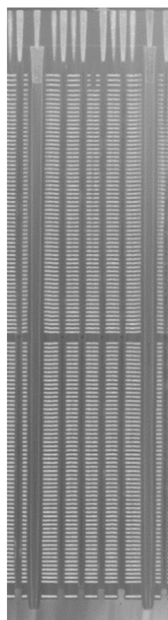
Figure: Cross section SEM of ON memory channel hole pattern after etching with novel chemistry cryogenic process.

Memory Technology

"Highly Scalable Metal Induced Lateral Crystallization (MILC) Techniques for Vertical Si Channel in Ultra-High (> 300 Layers) 3D Flash Memory" – Kioxia Corporation, Western Digital Corporation (Paper T7-1)

Kioxia and Western Digital demonstrate Metal-Induced Lateral Crystallization (MILC) techniques for fabricating 3D flash memory with ultra-high >300 layers. 14-μm-long macaroni silicon (Si) channels in vertical memory hole are fully single-crystallized by MILC. By using a newly developed nickel gettering technique, the 112 word-line-layered 3D flash memory exhibits >40% read noise reduction and 10x channel conductance without any degradation of cell reliability.





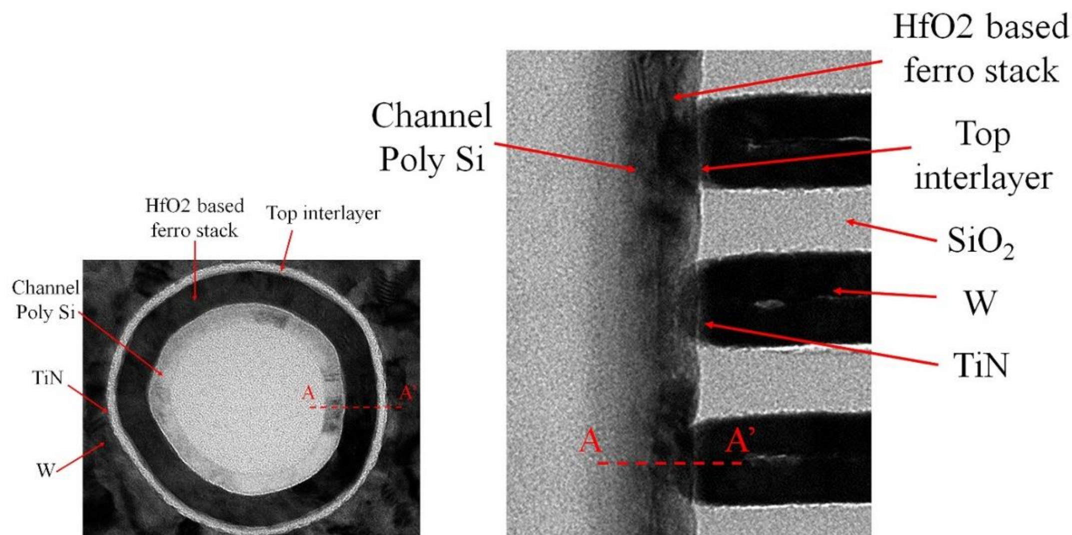
(a) (b)

Figures: TEM images of 3D flash memory with (a) over 300WL and (b) 112WL

Memory Technology

"QLC Programmable 3D Ferroelectric NAND Flash Memory by Memory Window Expansion using Cell Stack Engineering" – SK Hynix Inc. (Paper T7-2)

SK Hynix presents 3D ferroelectric NAND (Fe-NAND) Quad-level cell (QLC) operation using the 3D charge trap nitride (CTN) NAND test vehicle in mass production for the first time. They optimize the cell stack structure to improve the memory window (MW). The optimized top interlayer achieved the quadra-level-cell (QLC) with the minimum Vth gap margin of 0.24V. PE window expansion up to max 10.5V is realized for QLC programmable 3D Ferroelectric NAND Flash Memory.

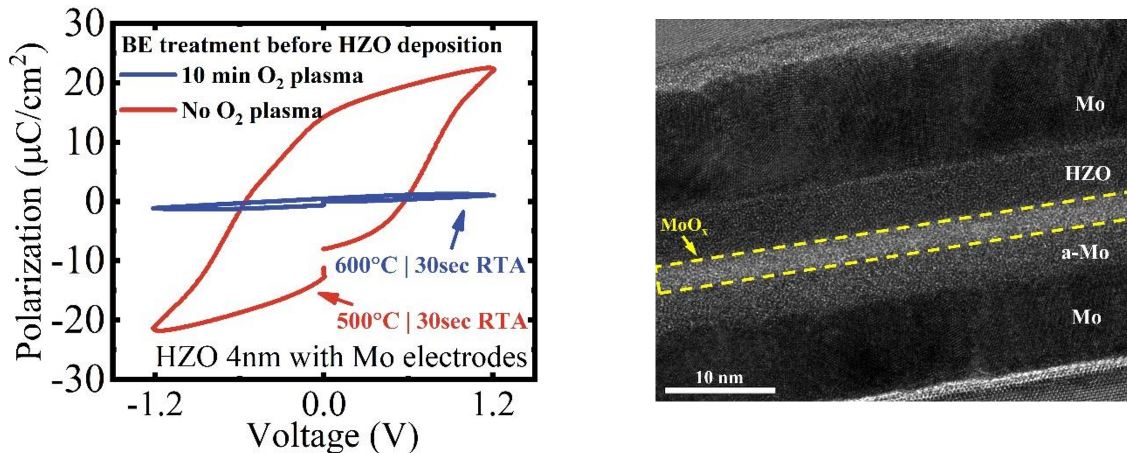


Figures: Fabricated 3D Fe-NAND (a) plan view, and (b) cross-sectional TEM images.

Advanced Process and Materials

"First Observation of Ultra-high Polarization (~ 108 $\mu C/cm^2$) in Nanometer Scaled High Performance Ferroelectric HZO Capacitors with Mo Electrodes" – Stanford University, Western Digital, University of Nebraska-Lincoln, University of Missouri, SLAC National Accelerator Laboratory (Paper T7-3)

A collaboration led by Stanford University demonstrates excellent ferroelectricity and endurance of 4nm-thick sub-100nm size $Hf_{0.5}Zr_{0.5}O_2$ (HZO) capacitors with Mo electrodes. They show (1) low crystallization temperature of 400℃, (2) low operation voltage of 1.2V with >10^{10} endurance cycles, (3) reduced wake-up effect and delayed fatigue by adding a CeO_2 stressor, (4) very large ferroelectricity of 108μC/cm^2 by carefully designed measurement system. This paper illustrates the importance of total material/process engineering and further potential of characteristics improvement in HfO_2-based ferroelectric capacitors.

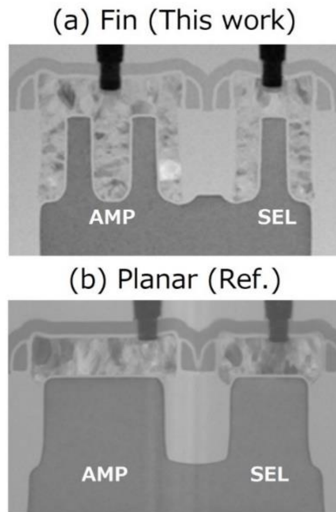


Figures: (Left) Ferroelectric P-V curves of HZO capacitors with TiN and Mo electrodes at different anneal conditions, (Right) top-down image of a fabricated sub-100nm size ferroelectric capacitor.

Image Sensor Technology

"Noise Performance Improvements of 2-Layer Transistor Pixel Stacked CMOS Image Sensor with Non-doped Pixel-FinFETs" – Sony Semiconductor Solutions, Sony Semiconductor Manufacturing Corporation (Paper T7-4)

Sony proposes a 2-Layer transistor pixel stacked CMOS image sensor with 2-fin non-doped pixel-fin FETs for the first time. Thanks to the non-doped channel and wider channel FinFET width, a 2.42x transconductance improvement, 15% random noise and 99.3% random telegraph signal reductions are reported.

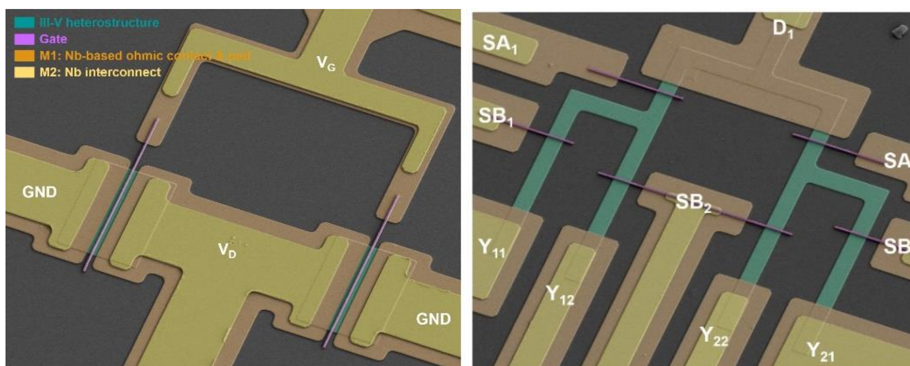


Figures: Cross sectional TEM images across the channel of (a) pixel-FinFET and (b) planar FETs of 2-Layer Transistor Pixel.

Beyond CMOS

"Cryogenic RF Transistors and Routing Circuits Based on 3D Stackable InGaAs HEMTs with Nb Superconductors for Large-Scale Quantum Signal Processing" – KAIST, KBSI, KNU, KANC (Paper T7-5)

Researchers at Korea Advanced Institute of Science and Technology, in collaboration with Korea Basic Science Institute, Kyungpook National University, and Korea Advanced Nano Fab Center, report on 3D stackable InGaAs HEMT-based cryogenec RF transistors and routing circuits integrated with Nb superconductors. The authors achieve a record high unity gain cutoff frequency of 601 GHz and unity power gain cutoff frequency of 593 GHz at 4 K with the smallest power dissipation among ever reported cryogenic RF transistors. Furthermore, using a novel structure with Nb superconductor and III-V heterostructure hybrid interconnect, the authors demonstrate routing circuits with 41% lower power dissipation compared to the conventional structure.



Figures: False colored SEM image of 3D stacked InGaAs HEMT-based (a) two-finger cryogenic RF transistor and (b) cryogenic 1-to-4 routing circuits with III-V heterostructure local interconnect and Nb superconductor global interconnect.

Circuits Highlights

Processors

"A 12-nm 0.62-1.61 mW Ultra-Low Power Digital CIM-based Deep-Learning System for End-to-End Always-on Vision" – MediaTek Inc. (Paper C3-4)

Researchers at MediaTek report on a Digital Compute-in-Memory (DCIM) promises an ultra-low power Deep Learning system for end-to-end Always-on-Vision. The authors present an SoC prototype comprised of a DCIM-based Deep Learning accelerator (DCIM-DLA), RISC-V microprocessor, and interface for off-chip image sensor connection. The DCIM supports mixed-precision computation to balance the power consumption and the desired accuracy. The prototype delivers a peak performance of 51.2 GOPS. Also, it achieves 57 TOPS/W energy efficiency and 85.7% accuracy with mixed-precision for human detection on MobileNet-V1. The power of the end-to-end system without an image sensor is 0.62 and 1.61 mW for 2 and 15 fps, respectively.

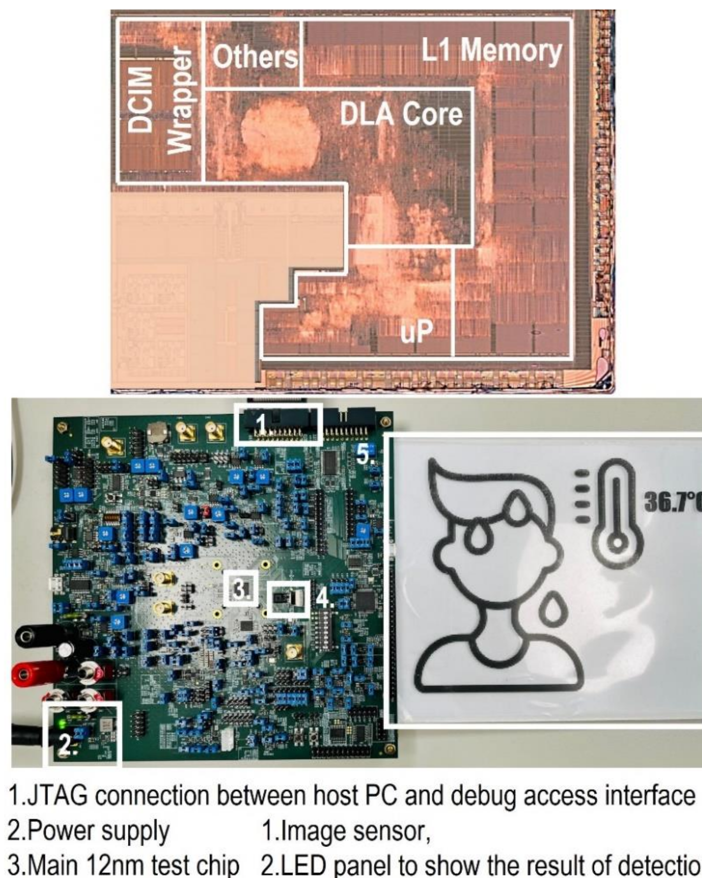


Figure: Presented is an SoC prototype comprised of a DCIM-base Deep Learning accelerator (DCIM-DLA), RISC-V microprocessor, and interface for off-chip image sensor connection. The power of the end-to-end system without an image sensor is 0.62 and 1.61 mW for 2 and 15 fps, respectively.

Imaging

"A 3.36 μm-pitch SPAD Photon-Counting Image Sensor Using Clustered Multi-cycle Clocked Recharging Technique with Intermediate Most-Significant-Bit Readout" – Sony Semiconductor Solutions Corp. (Paper C15-2)

Image sensors using SPAD pixels are expected to be able to capture images even in extremely dark situations because they produce images by capturing a single photon and directly counting its reaction numbers. On the other hand, the reaction number that a photon enters the SPAD pixel is counted, which increases the circuit size and power consumption in bright scenes. In this paper, the reset of the SPAD pixel is periodically controlled to suppress the SPAD pixel reaction in bright scenes, thereby reducing power consumption. The pixel size has also been shrunk by reducing the in-pixel counter bit to 8 bits because the upper digits are calculated by counting the number of changes in the most significant digit (MSB) of the in-pixel counter. The 22nm node enables the world's minimum pixel size of 3.36μm^2.

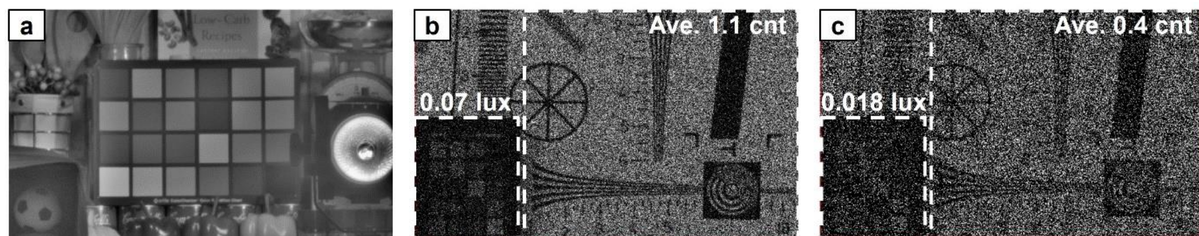


Figure: SPAD pixel allows for extremely dark scenes to be captured while still allowing bright scenes to be captured without saturation.

3D-Flash Memory

"A 1Tb 3b/Cell 3D-Flash Memory of more than 17Gb/mm^2 bit Density with 3.2Gbps Interface and 205MB/s Program Throughput" – KIOXIA Corp. (Paper C2-1)

Kioxia reports more than 210 word-line layers 1Tb 3b/cell 3D-Flash Memory with over 17Gb/mm^2 bit density. Physical 8-plane architecture realizes low read latency of 40μs and high program throughput of 205MB/s. High interface speed of 3.2Gbps is accomplished by reducing DQ area in the X direction to 41%. Hybrid row address decoders (X-DEC) can deal with the wiring congestion issue caused by the new architecture, minimizing the read latency degradation. One-pulse-two-strobe technique reduces sensing time by 18% and contributes to the achievement of 205MB/s program throughput.

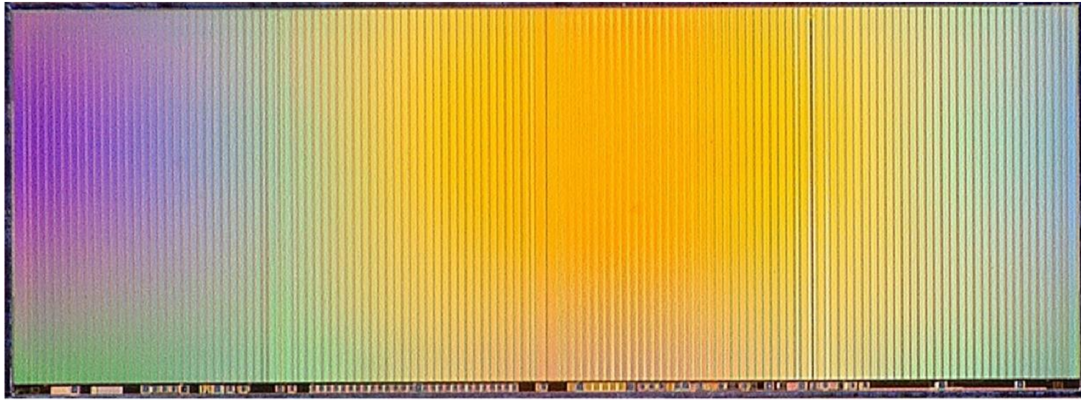


Figure. Die micrograph of 1Tb 3D-Flash memory.

SRAM Memory

"A 3-nm 27.6-Mbit/mm² Self-Timed SRAM Enabling 0.48 - 1.2 V Wide Operating Range with Far-End Pre-Charge and Weak-Bit Tracking" – TSMC Design Technology Japan (Paper C9-5)

TSMC presents a high-energy efficient cache SRAM on 3nm Fin-FET technology. High-performance computing systems place high demands on power efficiency. DVFS is widely used in recent designs to improve the power efficiency. In such DVFS system, on-dice cache memories are strongly required both high-speed operation at high over-drive voltage and green power operation at very low voltage. Two new DTCOs are introduced. One is the far-end bitline pre-charge circuit and the other is the weak-bit tracking circuit to support wide voltage range with fine grained DVFS. Introduced DTCOs improve performances against the 1) increase in wiring resistance and 2) increase in voltage-dependent sensitivity of transistor characteristics, which are challenges in cutting edge technologies. A test chip is fabricated in 3nm Fin-FET technology and demonstrates a high density of 27.6Mbit/mm^2 and 550MHz to 2.8GHz operation at 0.48V to 1.2V wide voltage range, achieved the best FoM (= Density x Fmax/VDD) among previous reports.

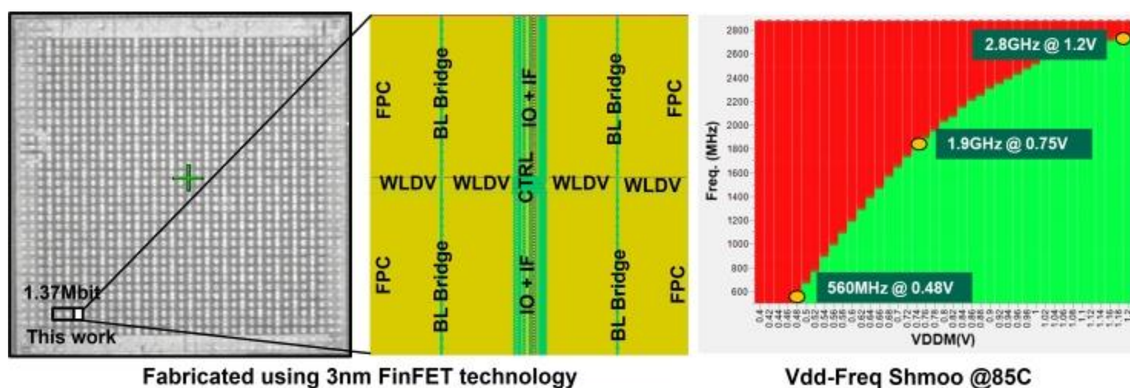
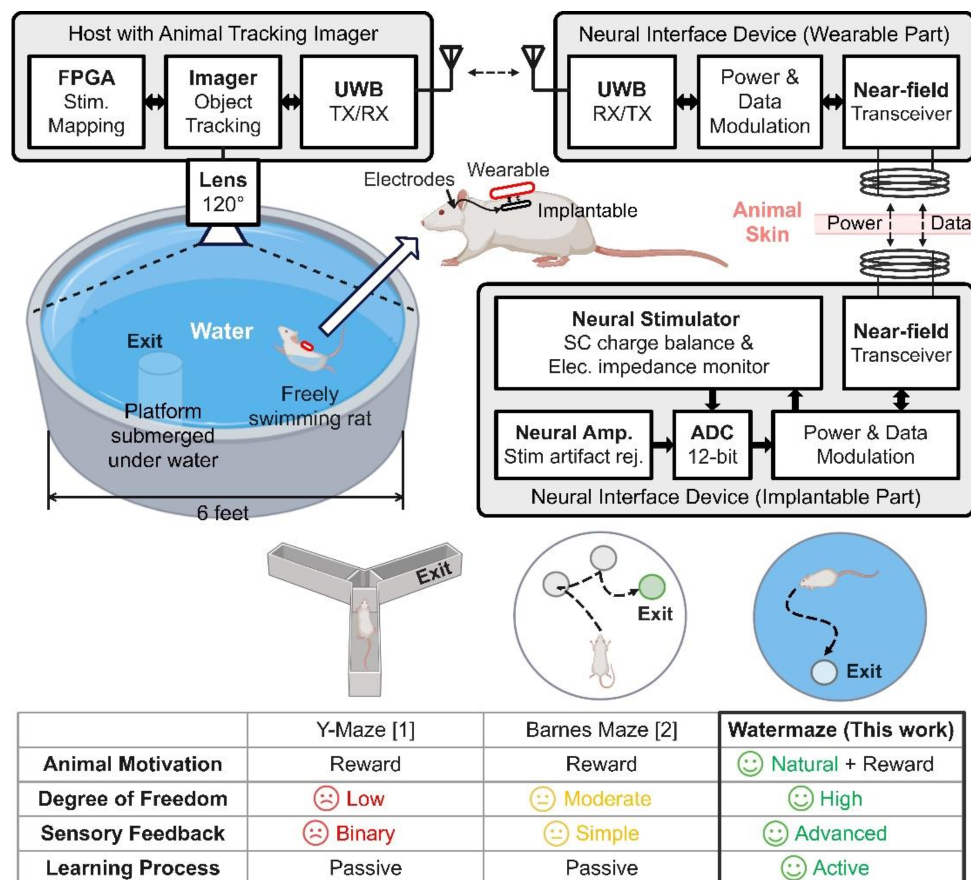


Figure: Die photograph of the fabricated 3nm Fin-FET test chip, layout plot of the 434kbit SRAM macro, and measured voltage-frequency shmoo plot.

Neural Interfaces

"A Wireless Sensor-Brain Interface System for Tracking and Guiding Animal Behaviors Through Goal-Directed Closed-loop Neuromodulation" – University of Toronto (Paper C1-1)

University of Toronto reports a wireless brain stimulation system capable of guiding a rodent in a water maze to the goal using real-time generated brain stimulation. The system is equipped with a FPGA-based controlling host with 160 x 160 image sensor, a wireless neural interface device consisting of fully-implantable and wearable parts. The host tracks a rat in the water maze and generate stimulation pattern for guiding. The authors successfully demonstrate that the rat can reach the submerged goal in the water maze with the help of the brain stimulation as fast in control experiments in which the goal can be visually seen.



	Y-Maze [1]	Barnes Maze [2]	Watermaze (This work)
Animal Motivation	Reward	Reward	Natural + Reward
Degree of Freedom	Low	Moderate	High
Sensory Feedback	Binary	Simple	Advanced
Learning Process	Passive	Passive	Active

Figure: Block diagram of the sensor-brain interface system and validity to use the water maze for demonstration.

Biomedical Circuits

"Wireless Body-Area Network Transceiver ICs with Concurrent Body-Coupled Powering and Communication using Single Electrode" – Southern University of Science and Technology (Paper C8-1)

Authors develop a BAN transceiver IC that provides power and data communication to each sensor node with a single electrode for continuous monitoring of health care information with

a body-mounted base station and multiple sensor nodes. When power transmission and communication are concurrently active, interference from the power transmission circuit on the base station saturates the receiver circuit, thus a self-interference cancellation circuit with suppression performance of more than 40dB was implemented to solve the problem. The sensor node circuit enables stable power reception by separating the ground of the power/data reception part and the data transmission part.

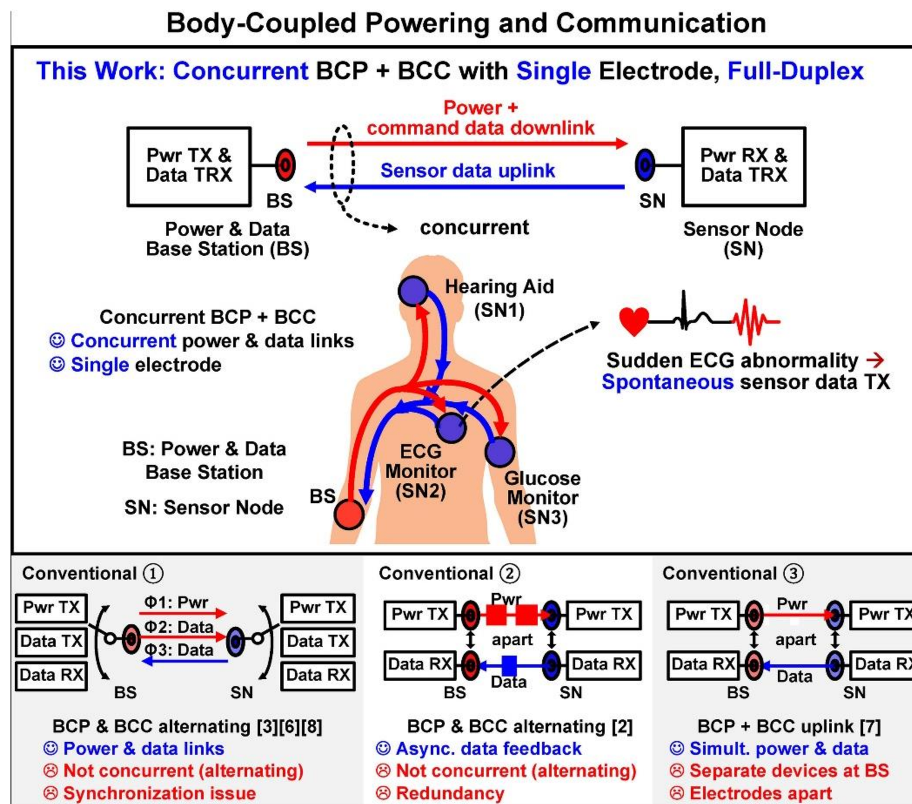


Figure. Concept of concurrent power transmission and data communication via human body using single electrode

Biomedical Circuits

"A Fingertip-Mimicking 12x16 200 um-Resolution e-skin Taxel Readout Chip with per-Taxel Spiking Readout and Embedded Receptive Field Processing" – KU Leuven (Paper C8-2)

KU Leuven reports an electronic skin (e-skin) taxel readout chip in 0.18µm CMOS technology, achieving the highest reported spatial resolution of 200µm, comparable to human fingertips. A key innovation is the integration on chip of a 12×16 taxel array with per-taxel signal conditioning frontend and spiking readout combined with embedded neuromorphic first-order processing through Complex Receptive Fields (CRFs). Compared to prior e-skin art, this work achieves around 100-7000× reduction in the system power consumption and >5 orders of magnitude reduction in the per-taxel power consumption, while enhancing the spatial resolution by 5× and doubling the sensor count. Experimental results show that Spiking Neural Network (SNN)-based classification of the chip's spatiotemporal spiking output for input tactile

stimuli, such as texture and flutter frequency, achieves excellent accuracies up to 97.1% and 99.2% of classification accuracy, respectively.

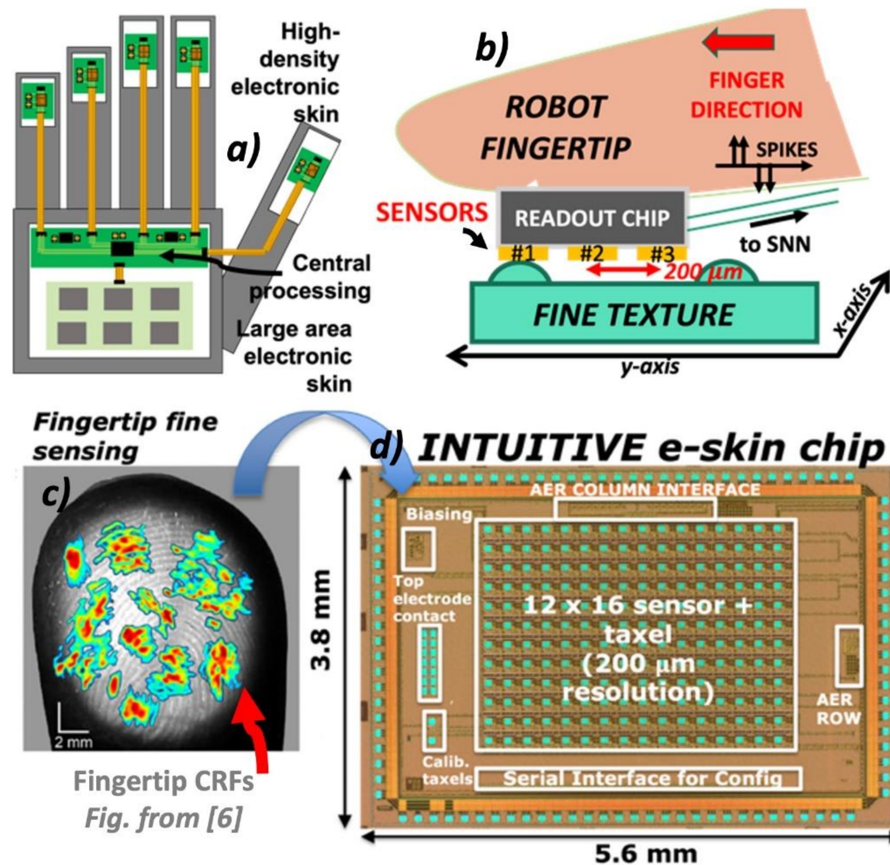


Figure: E-skin chip that mimics the tactile senses of human fingertips. The chip is mounted on the palm and fingertips of robot hands.

Digital Circuits

"Arvon: A Heterogeneous SiP Integrating a 14nm FPGA and Two 22nm 1.8TFLOPS/W DSPs with 1.7Tbps/mm2 AIB 2.0 Interface to Provide Versatile Workload Acceleration" – University of Michigan (Paper C7-1)

Researchers at the University of Michigan, in collaboration with Intel, report a heterogeneous system in a package (SiP) integrating a 14nm FPGA chiplet with two 22nm DSP chiplets through Embedded Multi-die Interconnect Bridges (EMIBs). The chiplets communicate via an Advanced Interface Bus (AIB) 1.0 interface and an AIB 2.0 interface. The SiP demonstrates the first-ever AIB 2.0 I/O prototype using 36µm-pitch micro bumps, achieving 4Gbps/pin at 0.10pJ/b (0.46pJ/b including adapter). The SiP is programmable, supporting workloads from neural network (NN) to communication processing (comm) and providing a peak performance of 4.14TFLOPS (FP16, half-precision floating-point). A compilation flow is developed to map workloads across FPGA and DSPs to optimize performance and utilization.

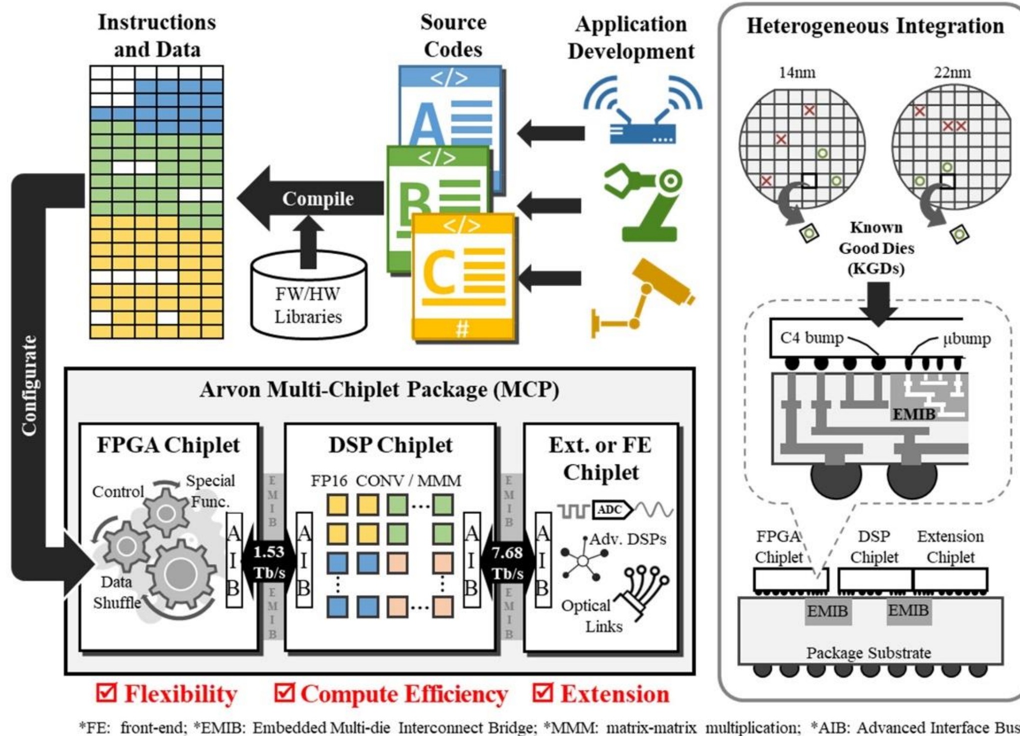


Figure: Arvon SiP heterogeneously integrating FPGA, DSP, FE chiplets for flexible workload mapping.

Wireline Receivers

"A 256 Gbps Heterogeneously Integrated Silicon Photonic Microring-based DWDM Receiver Suitable for In-Package Optical I/O" – Intel Corp. (Paper C6-2)

Intel proposes heterogeneously integrated silicon photonic microring-based dense wavelength division multiplexing (DWDM) receiver. A dither-based thermal control unit tunes micro-ring resonators in the optical demux to align with the laser grid with sub-pm resolution. The transceiver is implemented as a 28nm CMOS electric IC stacked on a silicon photonic IC. It achieves BER<1e-12 at 256 Gbps, 3.6 dBm optical power, 3.8 pJ/b energy efficiency by using uniform 200GHz spaced 8 wavelengths.

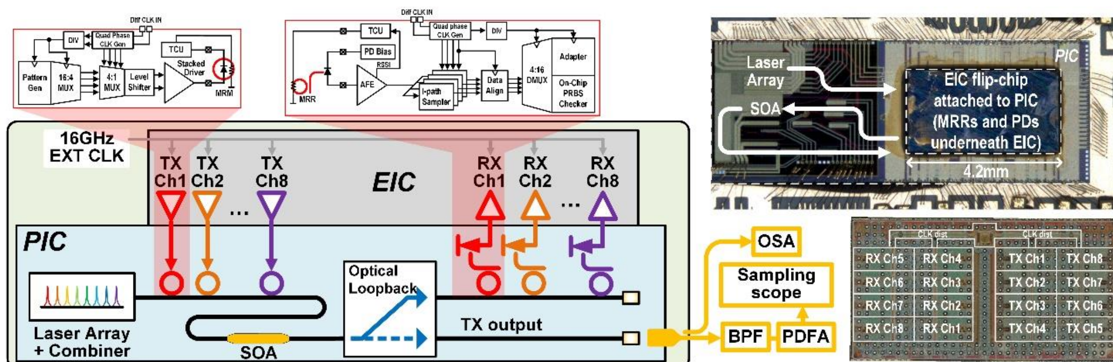
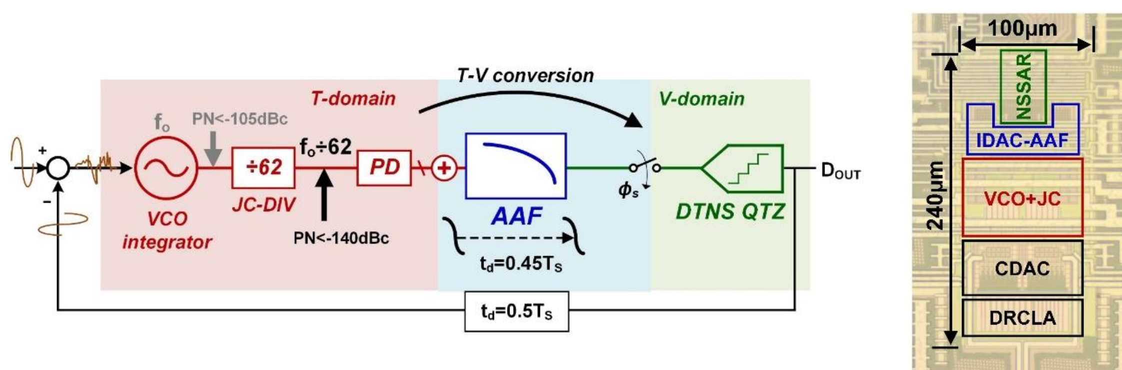


Figure: Heterogeneously integrated dense wavelength division multiplexing (DWDM) transceiver with electric IC and photonic IC detail. Photographs of transceiver assembly and electric IC.

Analog-to-Digital Converters

"A 0.024mm^2 84.2dB-SNDR 1MHz-BW 3rd-Order VCO-Based CTDSM with NS-SAR Quantizer (NSQ VCO CTDSM)" – University of Michigan (Paper C4-2)

University of Michigan proposes a new hybrid ADC architecture that uses a VCO-based continuous-time delta-sigma modulator (DSM) with a noise-shaping (NS) SAR quantizer. An anti-aliasing filter that bridges the VCO frontend with the NS SAR enables the time-domain information to be directly sampled as the voltage-domain information. The 28nm CMOS prototype achieves 84.2dB SNDR and 86.8dB DR within a 1MHz bandwidth while consuming 1.62mW at 100MS/s.



Figures: A new hybrid ADC architecture with an anti-aliasing filter that bridges the VCO frontend with the NS SAR and a chip photo of the 28nm CMOS prototype.

Analog Techniques

"An Energy-Efficient Impedance-Boosted Discrete-Time Amplifier Achieving 0.34 Noise Efficiency Factor and 389 MΩ Input Impedance" – ETH Zurich (Paper C19-2)

Researchers at ETH Zurich present a noise-efficient analog front end (AFE) for low-power sensor systems. The proposed AFE employs a low-noise amplifier based on series-parallel converters whose input impedance is boosted to 389MΩ using an input-resistance boosting loop and a capacitive positive feedback loop. This impedance-boosting technique provides 39x improvement compared to prior work. The AFE achieves the lowest reported noise efficiency factor and power efficiency factor of 0.34 and 0.1, respectively, while consuming 370nW power.

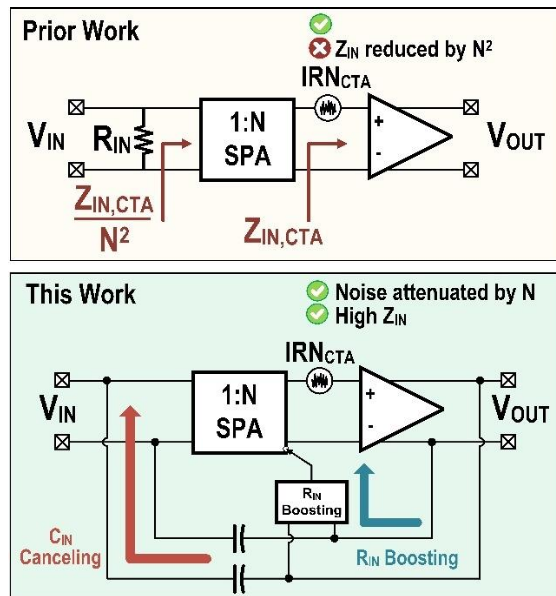


Figure: Input impedance-boosted analog front-end and comparison with prior work.

Frequency Generator

"A 122fsrms-Jitter and -60dBc-Reference-Spur 12.24GHz MDLL with a 102-Multiplication Factor Using a Power-Gating Technique" – Korea Advanced Institute of Science and Technology (KAIST) (Paper C26-5)

KAIST proposes a clock multiplier with low jitter and 12.24GHz output. A ring-oscillator-type multiplier is employed for saving layout area and achieves 0.066mm^2. With the conventional ring-oscillator-type multiplier, it is difficult to increase the output frequency but the proposed gating technique achieves a higher frequency output. In addition, the built-in calibration circuit can reduce reference spur to -60 dBc.

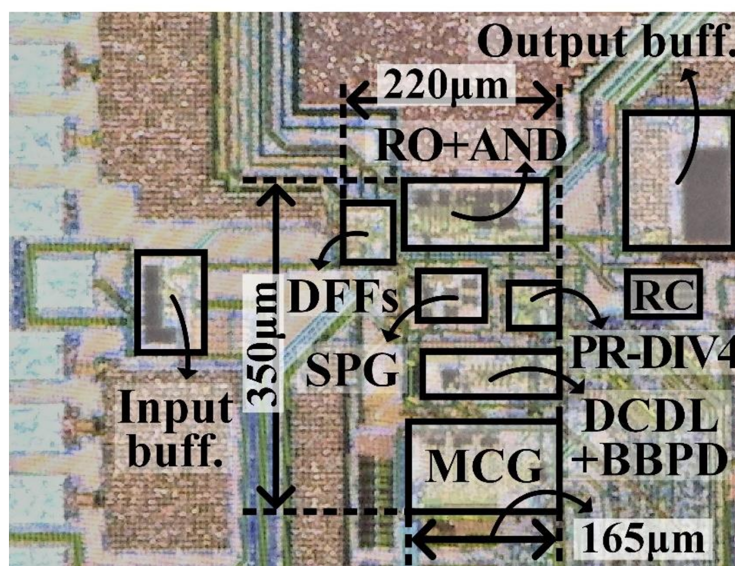


Figure: Low-jitter and 12.24GHz-output compact clock multiplier