Technical Highlights from the 2020 Symposia on VLSI Technology & Circuits

The 2020 Symposia on VLSI Technology & Circuits is a premiere international conference that defines the pace, progress and evolution of microelectronics, scheduled from June 15-18, 2020. The two Symposia will be held using a virtual format with a combination of live paper sessions as well as pre-recorded material.

The Symposia's overall theme, **"The Next 40 Years of VLSI for Ubiquitous Intelligence,"** integrates advanced technology developments, innovative circuit design, and the applications that they enable as part of our global society's transition to a new era of smart, connected devices and systems that change the way humans interact with each other.

Following are the Technology & Circuits highlighted papers that address this theme:

5G Transceivers & Silicon Photonic Switches

28GHz Dual-Polarized MIMO Beamformer

Researchers from Tokyo Institute of Technology, in cooperation with co-authors from NEC Corporation, will demonstrate a dual-polarized multiple-in, multiple-out (DP-MIMO) beamformer CMOS chip which supports two simultaneous 5G New Radio (NR) orthogonal frequency-division multiplexing (OFDMA)-mode streams operating at 256QAM. They constructed a 64-element phased array using 16 packaged beamformer chips (each element consisting of both vertical and horizontal polarized signals). Measurements using this DP-MIMO phased array with cross-polarization cancellation enabled an improvement of TX-to-RX EVM from 7.6% to 3.2%.

Paper C-F2-2 "A 28-GHz CMOS Phased-Array Beamformer Supporting Dual-Polarized MIMO with Cross-Polarization Leakage Cancellation" Jian Pang, et al., Tokyo Institute of Technology & NEC Corporation

5G Multi-standard RF Transceiver with Full Digital Interface in 14nm CMOS

Samsung will present a 5G NR FR1 RF Transceiver with legacy support implemented on a 14nm FinFET CMOS process. In contrast to many existing transceiver architectures, this design utilizes digitally-intensive circuit techniques as well as employing a fully-digital interface to a companion baseband modem chip. By utilizing this digital interface, the authors claim that the interface to the modem can be greatly simplified, reducing the number of modem downlink and uplink lines from 40 and 10 to 12 and 6, respectively.

Paper JFS2-6 "An RF Transceiver with Full Digital Interface Supporting 5G New Radio FR1 with 3.84Gbps DL/1.92Gbps UL and Dual-Band GNSS in 14nm FinFET CMOS" Sangwook Han, et al., Samsung Electronics

First Fully Packaged Silicon Photonics 8x8 Switch
Authors from IBM will demonstrate a silicon photonics-based network switch monolithically integrated with switching and control electronics. This new optical-based circuit switching technology enables fast switch reconfiguration times of less than 15ns while avoiding the high power of more conventional packet based electronic switches which required optical-to-electronic domain conversion. The authors claim that this technology uses a scalable process with simple flip-chip packaging.
Paper JFS1-3 "A Monolithically Integrated Silicon Photonics 8×8 Switch in 90nm SOI CMOS" Jonathan E. Proesel, et al., IBM

Optically Sampled ADC in 3D Integrated Silicon Photonics
The accuracy of conventional ADCs for high-frequency input signals is mainly limited by the sampling clock jitter. A group of researchers from UC Berkeley, MIT, LBNL, and CNSE are able to address this issue by implementing an ADC that uses low-jitter (<26 fs_{rms}) optical pulses to sample the input signal. A prototype two-channel ADC is realized in a 3D integrated platform with 65nm CMOS and silicon-photonics connected using high-density through-oxide-vias (TOVs). With optical pulses spaced at 250ps (4GS/s effective sampling rate), the ADC achieves SNDR of 40dB near DC and 37dB at 45GHz input.
Paper T-HL-3 "An Optically Sampled ADC in 3D Integrated Silicon-Photonics/65nm CMOS", N. Mehta et al., UC Berkeley/MIT/LBNL/CNSE

Memory Technology

Fast Thermal Quenching on Ferroelectric Thin Film with Record Polarization
Researchers from Hanyang University report an Al-doped HfO_2 (Al:HfO_2) ferroelectric film with record remnant polarization (Pr) and coercive electric field (Ec), demonstrating $2Pr = 100\mu C/cm^2$ and 2Ec= ~9.5MV/cm. They achieve this favorable ferroelectric (FE) characteristics by implementing fast quenching in DI water. The improvements are attributed to inducing higher stress/strain within Al:HfO2 thin film, leading to a stable orthorhombic phase (o-phase). Program/erase up to 10^6 cycles and 10 years retention characteristics are also obtained, potentially suitable for flash memory application.
Late-news Paper T-F2-5 "Fast Thermal Quenching on the Ferroelectric Al:HfO_2 Thin Film with Record Polarization Density and Flash Memory Application" B. Ku et al., Hanyang University

1.8Gb/s/pin 16Tb NAND Flash Memory Enabled by I/O Buffer
Authors from Samsung will reveal the technical details of their 3rd generation F-Chip, which enables their NAND Flash Memory stack to meet the performance requirements needed for peripheral component interconnect express (PCIe) Gen 4 host interfaces. The F-Chip, in combination with NAND achieves 1.8Gb/s/pin which is a 35% improvement from previous generation. This performance is enabled by built-in self test features which optimize sample timing and duty cycle error.
Paper JFS5-3 "A 1.8 Gb/s/pin 16Tb NAND Flash Memory Multi-chip Package with F-Chip of Toggle 4.0 Specification for High performance and High capacity Storage Systems" Jang-woo Lee, et al., Samsung Electronics

Extremely Scaled Hemi-Cylindrical 3D NAND Device
Macronix International reports an extremely scaled Hemi-Cylindrical (HC) 3D NAND device with large memory window. The proposed cell area ($0.009\mu m^2$/layer) is only ~32% of the standard GAA 3D NAND cell area, while it can produce extremely large >10V V_t memory window with excellent 100K endurance. The team at Macronix studies the size effect of HC device and finds that the larger (taller) HC devices may easily suffer parasitic edge leakage effect that causes programming saturation issue. A "wake-up" effect by an initial strong –FN erasing can introduce gate-injected electrons that electrically suppress the parasitic edge and, in turn, "wake-up" the device to produce a larger programming window. On the other hand, the smaller HC devices already show excellent memory windows without the need of wake-up. Good post-cycled retention and RTN performance are demonstrated for an extremely scaled "hero" HC device.
Paper T-M1-1 "An Extremely Scaled Hemi-Cylindrical (HC) 3D NAND Device with Large V_t Memory Window (>10V) and Excellent 100K Endurance" P-Y. Du et al., Macronix International

Quad Interface p-MTJ for STT-MRAM
Researchers led by Tohoku University have fabricated quad-interface perpendicular MTJ (Quad-MTJ) down to 33nm with their physical vapor deposition (PVD), reactive ion etching (RIE), and damage control integration process technologies on 300mm Si wafers. The 33nm Quad-MTJ technology achieves excellent endurance greater than 10^{11} write cycles due to higher write efficiency and low damage integration process technology. The results show that the Quad-MTJ technology is a promising way for low power, high speed, and enough reliable STT-MRAM with excellent scalability down to 1Xnm node.
Paper T-M3-1 "Scalability of Quad Interface p-MTJ for 1Xnm STT-MRAM with 10ns Low Power Write Operation, 10-years Retention and Endurance >10^{11}" S. Miura et al., Tohoku University

Digital Circuits, Hardware Security

Proactive Management of Clock Frequency Based on Predicted Power to Boost DSP Performance
Researchers from Qualcomm, in cooperation with University of Texas, will disclose a 7nm DSP which anticipates and mitigates supply voltage droops based on key microarchitectural events combined with a model of the power delivery network. Using this proactive clock-gating system (PCGS) to manage the clock frequency based on predicted power surges, the authors observed a benefit of 10% higher clock frequency or a corresponding reduction in supply voltage by reducing V_{min} by 5%.
Paper C-2-1 "A Proactive Voltage-Droop-Mitigation System in a 7nm Hexagon™ Processor" Vijay Kiran Kalyanam, et al., Qualcomm Technologies, Inc. & University of Texas

Side Channel Attack Resistant Crypto-Processor
Intel researchers will unveil a Rivest–Shamir–Adleman (RSA)-4K crypto-processor designed to be resistant to side channel attacks (SCA). The approach revealed in this paper is intended for post-quantum RSA cryptosystems to reduce the viability of brute-force attacks using quantum computers. This public key encryption approach utilizes a combination of circuit obfuscation and randomization which limits SCA-resistant circuit area/performance overhead to 3%.

Paper C-1-4 "A 435MHz, 2.5Mbps/W Side-Channel-Attack Resistant Crypto-Processor for Secure RSA-4K Public-Key Encryption in 14nm CMOS" Raghavan Kumar, et al., Intel Corporation

Artificial Intelligence, Machine Learning, and Quantum Computing

32GHz Cryogenic Superconducting 4-bit Processor

Researchers from Kyushu University and Nagoya University have developed a Single-Flux-Quantum (SFQ) 4-bit processor operating at 32GHz in the first of its kind demonstration using a gate-level pipelined architecture. To achieve the superconductor SFQ properties, the experiment was cooled using liquid helium at 4.2K. Unlike conventional common clock pipelined design used in commercial processors, this design used a point-to-point clock synchronization in which the clock flows in the same or opposite direction of the data, depending on whether the logic is feedforward- or feedback-based, respectively. The authors claim 2.5 tera-operations per watt for the 4-bit processor implementation.

Paper C-A3-5 "32 GHz 6.5 mW Gate-Level-Pipelined 4-bit Processor using Superconductor Single-Flux-Quantum Logic" Koki Ishida, et al., Kyushu University & Nagoya University

28nm FD-SOI technology at cryogenic temperatures for quantum computing

Quantum computers need high performance and low power control electronics with operating temperatures ranging from 4.2K low temperature (LT) down to well below 1K. Researchers at CEA-LETI and STMicroelectronics propose FD-SOI transistors as the CMOS technology solution for this application, and report their findings in evaluating 28nm FD-SOI transistors for variability and performance at these ultra-low temperatures. High performance is achieved at ultralow temperature (ULT) for short channel transistors, with I_{ON}>1mA/µm and I_{OFF} below the equipment accuracy <1fA, in particular by keeping advantage of forward back biasing (FBB), with the same efficiency from room temperature (RT) down to 100mK. The physical origins of MOSFET mismatch at ULT are studied, highlighting the impact of charge fluctuation increase on both threshold voltage (V_{TH}) and current gain factor (β) variabilities. They also demonstrate that the increase of V_{TH} and β variabilities at low temperature remains reasonably low in comparison to RT values and other CMOS technologies, so that it should not be detrimental to circuit operation in this range of temperatures.

Paper T-N3-1 "Variability Evaluation of 28nm FD-SOI Technology at Cryogenic Temperatures down to 100mK for Quantum Computing" B. Cardoso Paz et al., CEA-Leti/STMicroelectronics/Institut Néel

High compute utilization hardware for AI training & inference

IBM will demonstrate a processor core which can be applied to both AI training as well as inference applications. The authors of this paper developed innovations such as heterogeneous compute engines with architectural programmability and dataflow mappings configured to support a broad set of AI workloads. Combined with a software-controlled network interface, this resulted in improvements in hardware utilization which led to notable improvements in training area efficiency (0.30TFLOPS/mm^2 at 0.62V) and training compute efficiency (1.4TFLOPS/W at 0.54V).

Paper C-A1-1 "A 3.0 TFLOPS 0.62V Scalable Processor Core for High Compute Utilization AI Training and Inference" Jinwook Oh, et al., IBM

Monolithic 3D Integration of RRAM Array
Researchers at the University of Tokyo have successfully integrated RRAM arrays with indium gallium zinc oxide (IGZO) access transistors monolithically in a 3D stack, and demonstrated basic functionality of in-memory computing in the 3D neural net. They have achieved uniform memory characteristics of 1T1R cells at each layer, and demonstrated XNOR for binary neural network AI applications for the first time. The impact of RRAM bit error rate on neural networks is also investigated, and they conclude that 3D neural networks built by this architecture have high potential to enable area-efficient, low power and low-latency computing.
Paper T-HL-4 "A Monolithic 3D Integration of RRAM Array with Oxide Semiconductor FET for In-memory Computing in Quantized Neural Network AI Applications" J. Wu et al., The University of Tokyo

SOC/3D Packaging

5G and AI Integrated High Performance Mobile SoC with 7nm EUV FinFETs
Qualcomm Technologies discloses its Snapdragon™ 765 mobile platform and world's first integrated 5G platform supporting both mmWave and sub-6GHz applications using industry-leading 7nm EUV FinFET technology. Snapdragon 765 is intended to unite 5G and AI to power select premium tier experiences on a global scale, and it exhibits 20% improvement in performance and 35% lower power consumption over its predecessor Snapdragon 730 (8nm FinFET). Key features include their new technology integration feature, mixed diffusion break (MDB), to boost device performance boost, dual poly pitch process to enable power-perf efficient design architecture, and process design co-development to achieve low voltage logic/memory operation. Further process-design co-optimization also reduces CPU V_{min} by 80mV, enabling a premium tier performance experience with integrated 5G and AI mobile SOC platforms.
Paper T-HL-1 "5G and AI Integrated High Performance Mobile SoC Process-Design Co-Development and Production with 7nm EUV FinFET Technology" J. Deng et al., Qualcomm Technologies/ Samsung Electronics

Low Temp SoICTM Bonding & Stacking for High Bandwidth Memory
TSMC researchers ran a low temperature bonding and stacking technology to realize a 12-high (12-Hi) and 16-high (16-Hi) die stack for high bandwidth memory (HBM) applications. The daisy chains in the 12-Hi structure, incorporating more than ten thousand TSVs and bonds were tested, and linear I-V curves were obtained, demonstrating good bonding and stacking quality. The electrical link from a base logic die to top DRAM is built up to study the bandwidth and power consumption. Compared to μbump technology, the bandwidth for 12-Hi and 16-Hi structure using the SoIC technology shows improvement of 18% and 20%, respectively, and the power efficiency demonstrates the improvement of 8% and 15%, respectively. Also, the thermal performance for the 12-Hi and 16-Hi SoIC-bond structures are improved by 7% and 8%, respectively. The authors report the scalability projection of bonding pitch to increase pin counts and die thickness to reduce the TSV parasitic capacitance for future HBM applications.
Paper T-H1-1 "Low Temperature SoICTM Bonding and Stacking Technology for 12/16-Hi High Bandwidth Memory (HBM)" C.H. Tsai, et al., Taiwan Semiconductor Manufacturing Company

Sensor and Display Circuits

Piezoresistive Pressure Sensor for Ultra Low-Power IoT Applications
Researchers affiliated with University of Michigan and CubeWorks will present a piezoresistive pressure sensor integrated in a fully encapsulated system with a total volume of 20mm^3. By utilizing a highly duty-cycled excitation of a Wheatstone Bridge based sensor followed by efficiency amplification and sub-ranging ADC sampling, the system achieves state-of-art energy efficiency of 6.1 nJ·mmHg^2.
Paper C-B3-2 "A Pressure Sensing System with ±0.75mmHg (3σ) Inaccuracy for Battery-Powered Low Power IoT applications" Seokhyeon Jeong, et al., University of Michigan & CubeWorks

LiDAR Sensor with Embedded Interference Filter
Researchers from Sungkyunkwan University, SOS LAB, Samsung, and UNIST will reveal a LiDAR sensor with an on-chip single-photon avalanche diode array. Using a time-of-flight detection approach, the LiDAR system beam scanner uses dual laser diodes which transmit two successive pulses shifted by a predetermined amount of time. This time shifting enables differentiation between the reflected signal and inter-LiDAR interference. To reduce area and complexity, the system includes an integrated mixed signal histogramming function that mitigates the need for complex averaging logic or memory arrays. Combined with a rotating polygon mirror and MEMS mirror for horizontal and vertical scanning, respectively, the LiDAR system achieves 120° x 8° field of view with a 48m distance range and accuracy of 11.68cm.
Paper C-B2-2 "A 36-channel SPAD-integrated Scanning LiDAR Sensor with Multi-event Histogramming TDC and Embedded Interference Filter" Hyeongseok Seo, et al., Sungkyunkwan University

Biomedical Circuits

Artificial Iris Contact Lens to Combat Human Eye Deficiencies
IMEC, with co-authors from Ghent University and KU Leuven, will detail a fully encapsulated, artificial iris embedded in a smart contact lens to mitigate problems associate with human eye iris deficiencies. The variable iris aperture is implemented using four concentric rings of varying diameter on embedded lens LCD. To operate for an entire day without having to charge the embedded power source, the system power consumption is 1.9mW to control the iris aperture as well as operate the eye blink sensor, light sensor, and corresponding compute logic. Daily charging occurs via an NFC receiver coil and IC after the lenses are removed for storage.
Paper C-B1-2 "An Artificial Iris ASIC with High Voltage Liquid Crystal Driver, 10nA Light Range Detector and 40nA Blink Detector for LCD Flicker Removal" Bogdan C. Raducanu, et al., imec

Advanced CMOS Technologies

Buried Power Rail Integration with Si FinFETs for CMOS Scaling beyond 5nm
Researchers at IMEC report their work on demonstrating the feasibility of buried power rail (BPR) with FinFET architecture and proposed BPR as a key scaling booster for CMOS extension beyond the 5nm node. Their paper demonstrates, for the first time, the integration of tungsten (W) BPR lines with Si FinFETs. They find characteristics of

CMOS in close proximity to floating BPR to be similar to the characteristics of CMOS without BPR. Moreover, the W-BPR interface with ruthenium (Ru) via contact can withstand more than 320h of electromigration (EM) stress at 4MA/cm2 and 330°C, making Ru a candidate for via metallization to achieve a low resistance contact strategy for BPR.

Paper T-HL-6 "Buried Power Rail Integration with Si FinFETs for CMOS Scaling beyond the 5nm Node" A. Gupta et al., imec

Seven-Level-Stacked Nanosheet GAA Transistors

Gate-all-around (GAA) nanosheet transistors offer higher performance thanks to higher effective channel width (W_{eff}) as compared to the state of the art FinFET transistors. Researchers at CEA-LETI-MINATEC evaluate the trade-off of increasing the W_{eff}/footprint to boost device performance with process complexity. They experimentally demonstrate, for the first time, GAA nanosheet (NS) transistors with seven-level stacked nanosheet transistors fabricated using a replacement metal gate process, inner spacer, and self-aligned contacts. They demonstrate excellent gate controllability with extremely high current drivability (3mA/μm at V_{DD}=1V) and a 3X improvement in drain current over usual two-level stacked NS GAA transistors.

Paper T-C1-2 "7-Level-Stacked Nanosheet GAA Transistors for High Performance Computing" S. Barraud et al., CEA-LETI-MINATEC

Air Spacer with Self-Aligned Contact & Contact Over Active Gate

Air spacer technology has long been recognized as the most effective approach to reduce parasitic capacitance. IBM researchers will report an improved air spacer that is successfully co-integrated on FinFET transistors with Self-Aligned Contacts (SAC) and Contacts Over Active Gate (COAG). The new integration scheme enables air spacer formation agnostic to the underlying transistor architecture, thus paving the way for a seamless adoption of air spacer in FinFET and Gate-All-Around (GAA) technologies. A reduction in effective capacitance (C_{eff}) by 15% is experimentally demonstrated. The power/performance benefits achieved by the new air spacer exceed the benefits of scaling FinFET from 7nm node to 5nm node.

Paper T-HL-5 "Improved Air Spacer Co-Integrated with Self-Aligned Contact (SAC) and Contact Over Active Gate (COAG) for Highly Scaled CMOS Technology" K. Cheng et al., IBM Research

Heterogeneous Integration, Non-Silicon Substrates/Materials & Devices

GaN & SiC Transistors Enabled By 3D Monolithic Integration

Intel expands their work in heterogeneous 3D GaN/Si integration by demonstrating both Si P- and NMOS FinFET transistors monolithically integrated with GaN transistors on 300mm Si(111) wafers. With the Si FinFET architecture, the researchers at Intel are able to take advantage of the fin orientations of the transferred Si(100) crystal to fabricate both high performance Si P- and NMOS transistors. Furthermore, they demonstrate a variety of GaN transistor innovations, including enhancement (e-mode) and depletion mode (d-mode) GaN NMOS transistor with high ID=1.8mA/μm; GaN Schottky gate transistor producing high saturated power of 20dBm with peak PAE=57% at 28GHz; high

performing, low leakage cascode and multi-gate GaN transistors; and GaN Schottky diodes with ultra-low C_{OFF} for ESD protection, all integrated on 300mm Si(111) wafer.
Paper T-HL-2 "GaN and Si Transistors on 300mm Si(111) enabled by 3D Monolithic Heterogeneous Integration" H.W. Then et al, Intel Corp

Vertical-Channel FET w/High-Thermal-Tolerance In-Al-Zn Oxide Channel
A research team at Kioxia have demonstrated, for the first time, a surrounding gate vertical-channel FET with gate length of 40nm by introducing backend-of-line (BEOL) process compatible novel oxide semiconductor (OS) In-Al-Zn-O as a channel material. Fabricated FETs exhibit high scalability by excellent thermal stability (~ 420°C) compared to conventional In-Ga-Zn-O-channel FETs, with high mobility (12.7cm^2/Vs) characteristics. Furthermore, the vertical-channel FET also exhibits excellent reliability and stable operation without floating body effect. Endurance of more than 10^{11} cycles is also demonstrated. This work opens a pathway to realization of high-performance BEOL transistor for 3D-LSI applications.
Paper T-H2-2 "Surrounding Gate Vertical-Channel FET with Gate Length of 40nm Using BEOL Compatible High-Thermal-Tolerance In-Al-Zn Oxide Channel", H. Fujiwara et al., Kioxia Corp