SEMICONDUCTORS | AI | DESIGN | PACKAGING | MEMS | DISPLAYS

SEMICONDUCTOR DIGEST

NEWS AND INDUSTRY TRENDS

JULY 2023

The State of the Semiconductor Industry 2023

www.semiconductordigest.com

Nikon

Big Ideas,
The Best Team,
The Right Solutions.

Nikon.
Technology to Do More.

The world has changed dramatically, and Nikon recognizes that technology must do more to serve society and our planet.

Our semiconductor lithography and inspection systems bring together the winning combination of over a century of Nikon innovation and the most dedicated employees in the world to deliver the right solutions for our customers.

www.nikonprecision.com

Visit us at SEMICON West 2023
July 11-13, 2023 - Booth #1028 South Hall

SEMICONDUCTOR DIGEST

NEWS AND INDUSTRY TRENDS

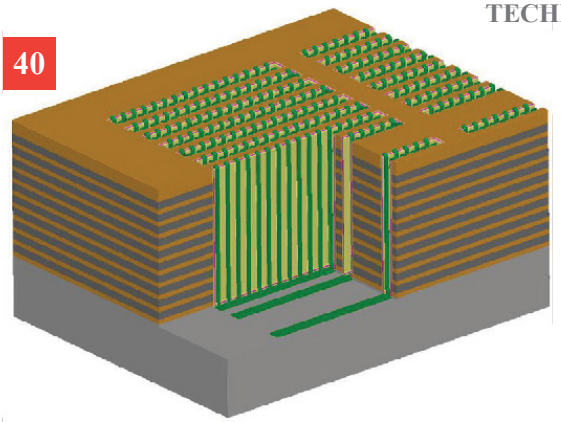
July 2023 | Volume 5 Number 5

COVER: ElevATE Semiconductor is pioneering next-generation solutions for semiconductor testing. Shown (left to right): David Kenyon: CEO, Simon Leigh: VP of Engineering, and Tim Bakken: VP of Global Sales.

Columns

Departments

FEATURES



We continuously strive to minimise the environmental impact of the semiconductor industry in our natural world and environment we live in now and for our future.

EDWARDS

Editorial

AI in Semiconductor Manufacturing

ARTIFICIAL INTELLIGENCE IS POISED TO IMPACT VIRTUALLY every aspect of our lives moving forward. It's also playing an increasingly important role in the manufacturing of semiconductors.

At the recent imec-hosted ITF World held in Antwerp, Belgium in May, Jensen Huang, president, CEO and Member of the Board of Directors at NVIDIA, gave some insights into how NVIDIA is working with companies such as TSMC, ASML, Applied Materials (AMAT), D2S, IMS Nano Fabrication and Synopsys to bring AI to chip making.

"The first wave of AI that focused on computer vision and speech recognition has achieved superhuman capabilities and has opened up multi-trillion dollar opportunities in robotics, autonomous vehicles and manufacturing," Huang said. "Advanced chip manufacturing requires over a thousand steps, producing features the size of a biomolecule. To make chips with quadrillions of features, each step must be nearly perfect to yield any output. Sophisticated computational sciences are performed at every stage to compute the features to be patterned and to do defect detection for inline process control. Chip manufacturing is an ideal application for NVIDIA accelerated computing and AI."

Huang said D2S and IMS Nano Fabrication build mask writers using e-beams to create photoresist patterns on a mask. "Nvidia GPUs do the pattern rendering and mask process correction," he said. TSMC and KLA use EUV and DUV illumination for mask inspection. "NVIDIA GPUs process classical physics modeling, and deep learning to generate synthetic reference images and detect defects," he said.

KLA and AMAT use NVIDIA for their e-beam wafer inspection systems and for their optical wafer inspection systems, which detect defects smaller than the pixel. "The data rate is extremely high at 40 gigapixels per second to achieve this as part of process control," Huang said.

TSMC, ASML, Synopsys and NVIDIA recently announced plans to accelerate computational lithography. "Computational lithography simulates Maxell's equations of light passing through optics and interacting with photoresists. It is a vital step in chip making and the largest computational workload in chip design and manufacturing, consuming tens of billions of CPU hours annually," Huang said, noting that massive data centers run 24/7 to create reticles for new chips. "We have already accelerated the processing by 50 times," he added.

Christophe Fouquet, CBO & Board of Management at ASML, who also spoke at ITF World said that, in the last few years, the company has shifted to AI and massive use of metrology. What does it mean? The number of parameters we use in our models to describe the litho process have increased by a factor of 800," he said. He said in ASML's scanners, more than 50,000 parameters can be tuned for every die exposed. "It's allowed us to improve our accuracy, alignment and overlay," he said.

AI is being used to make AI chips — the cobbler's children have some fine new shoes, indeed!

—Pete Singer, Editor-in-Chief

SEMICONDUCTOR DIGEST
NEWS AND INDUSTRY TRENDS

Kerry Hoffman, Publisher
kerryh@semiconductordigest.com
978-580-4205

Pete Singer, Editor-in-Chief
psinger@semiconductordigest.com
978-470-1806

Shannon Davis, Web and News Editor
sdavis@semiconductordigest.com

Dave Lammers, Contributing Editor
dplammers@att.net

John Blyler, Contributing Editor
jblyler@semiconductordigest.com

Dick James, Contributing Editor
dickjames@siliconics.ca

Cindy Chamberlin, Art Director

Rich Mehta, Website Design

July 2023 VOL. 5 NO. 5 •

Semiconductor Digest ©2023

ISSN 2643-7058 (print) ISSN 2643-7074 (online).
Semiconductor Digest is published eight times a year by Gold Flag Media LLC, 58 Summer St., Andover, MA 01810.

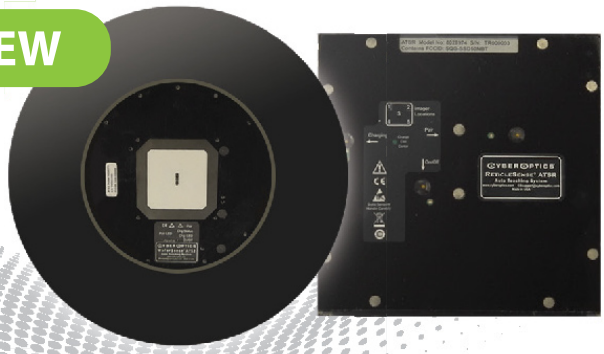
Copyright © 2023 by Gold Flag Media LLC.
All rights reserved. Printed in the U.S.

For subscription information, please visit www.semiconductordigest.com/subscribe.

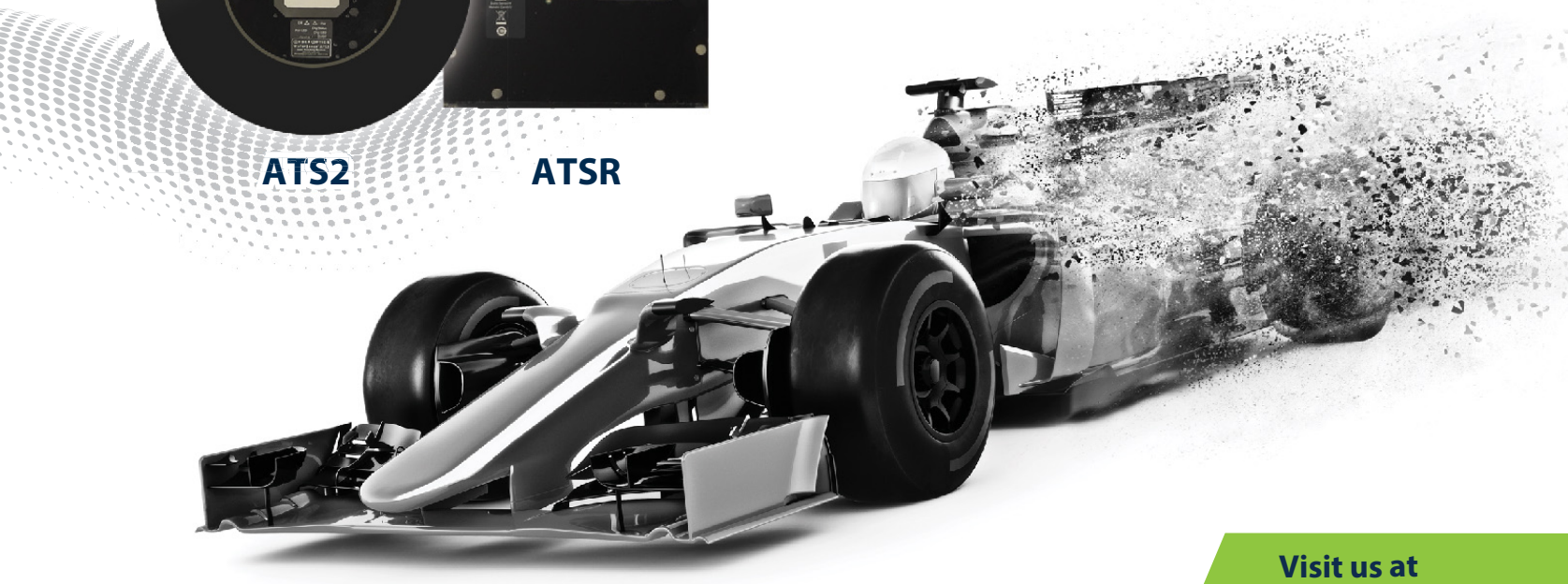
Speed is your friend.

Dramatically speed up semiconductor tool alignment.

NEW



ATS2 ATSR



Visit us at SEMICON West Booth # 964

Improve yields and tool up time while saving time and expense.

Introducing the next-generation, wireless WaferSense® and ReticleSense® Auto Teaching System™ (ATS2™/ATSR™) with CyberSpectrum™ software.

Signicantly faster multi camera system teaching for accurate wafer and reticle hand-off calibration, proper alignment, and set-up for semiconductor tools. ATS2 and ATSR "see" inside in real-time to capture three dimensional off-set data (x, y and z) to quickly teach wafer and reticle transfer positions - all without opening the tools.

Nordson TEST & INSPECTION

CYBEROPTICS®

www.cyberoptics.com

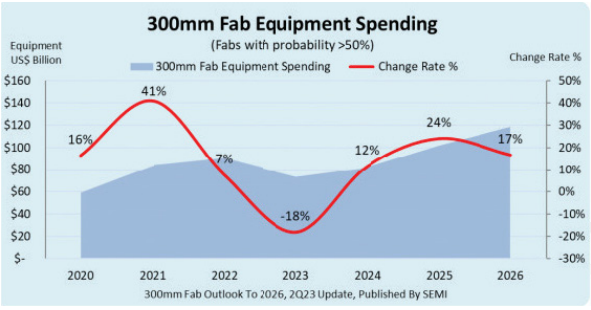
Copyright © Nordson Corporation. All rights reserved.

news

Global 300mm Fab Equipment Spending Forecast to Reach Record $119 Billion in 2026, SEMI Reports

Global 300mm fab equipment spending for front-end facilities next year is expected to begin a growth streak to a US$119 billion record high in 2026 following a decline in 2023, SEMI highlighted in its quarterly 300mm Fab Outlook Report to 2026. Strong demand for high-performance computing, automotive applications and improved demand for memory will fuel double-digit spending in equipment investments over the three-year period.

After the projected 18% drop to US$74 billion this year, global 300mm fab equipment spending is forecast to rise 12% to $US82 billion in 2024, 24% to US$101.9 billion in 2025 and 17% to US$118.8 billion in 2026.



"The projected equipment spending growth wave underscores the strong secular demand for semiconductors," said Ajit Manocha, SEMI President and CEO. "The foundry and memory sectors will figure prominently in this expansion, pointing to demand for chips across a wide breadth of end markets and applications."

Regional growth

Korea is expected to lead global 300mm fab equipment spending in 2026 with US$30.2 billion in investments, nearly doubling from US$15.7 billion in 2023. Taiwan is forecast to invest US$23.8 billion in 2026, up from US$22.4 billion this year, and China is projected to log US$16.1 billion in spending in 2026, an increase from US$14.9 billion in 2023. Americas equipment spending is expected to nearly double from US$9.6 billion this year to US$18.8 billion in 2026.

Segment growth

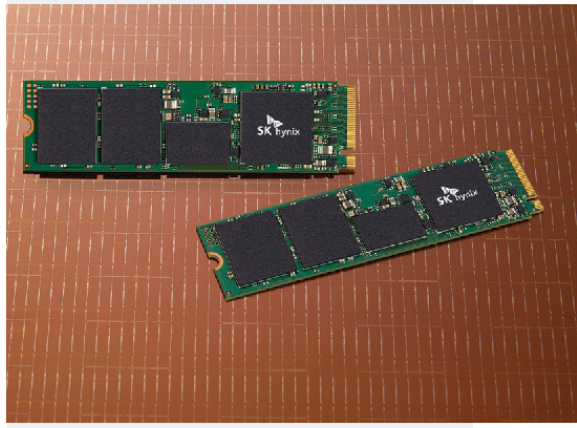
Foundry is projected to lead other segments in equipment spending at US$62.1 billion in 2026, an increase from US$44.6 billion in 2023, followed by memory at US$42.9 billion, a 170% increase from 2023. Analog spending is forecast to increase from US$5 billion this year to US$6.2 billion in 2026. The microprocessor/microcontroller, discrete (mainly power devices), and optoelectronics segments are expected to see spending declines in 2026, while investments in logic are forecast to rise.

The SEMI 300mm Fab Outlook Report To 2026 report lists 369 facilities and lines globally, including 53 high-probability facilities expected to start operation during the four years starting in 2023.

SK hynix Begins Mass Production of 238-Layer 4D NAND

SK hynix Inc. has started mass production of its 238-layer 4D NAND Flash memory, following the development in August 2022, and that product compatibility test with a global smartphone manufacturer is underway.

"SK hynix has developed solution products for smartphones and client SSDs which are used as PC storage devices, adopting the 238-layer



NAND technology, and has moved into mass production in May," the company said. "Given that the company secured world-class competitiveness in price, performance and quality for both 238-layer NAND and the previous generation 176-layer NAND, we expect these products to drive earnings improvement in the second half of the year."

The 238-layer product — the smallest NAND in size — has a 34% higher manufacturing efficiency compared to the previous generation of 176-layer, resulting in a significant improvement in cost competitiveness.

Besides, with a data-transfer speed

Continued on page 10

YOUR COMPLETE SUBFAB PARTNER.

Vacuum pumps.
Gas abatement.
Leak detection.
Global service.

VISIT US AT SEMICON WEST BOOTH# 541

ask about our **LEAK TEST SOLUTIONS**

From load lock pump down to the harshest byproduct processes, Busch provides leading semiconductor Fabs with vacuum pumps and abatement solutions, full SubFab management, and support tailored to your needs.

Contact us today!
Scan to learn more

1-800-USA-PUMP | info@buschusa.com | **buschusa.com**

BUSCH
VACUUM SOLUTIONS

Global Semiconductor Materials Market Revenue Reaches Record $73 Billion in 2022

Global semiconductor materials market revenue grew 8.9% to $72.7 billion in 2022, surpassing the previous market high of $66.8 billion set in 2021, SEMI, the global industry association representing the electronics manufacturing and design supply chain, reported in its Materials Market Data Subscription (MMDS).

Wafer fabrication materials and packaging materials revenue in 2022 reached $44.7 billion and $28.0 billion, respectively, increasing 10.5% and 6.3%. The silicon, electronic gases, and photomask segments showed the strongest growth in the wafer fabrication materials market, while the organic substrates segment largely drove packaging materials market growth.

For the 13th consecutive year, Taiwan, at $20.1 billion, was the world's largest consumer of semiconductor materials on the strength of its foundry capacity and advanced packaging base. China continued to register strong year-over-year results, ranking second in 2022, while Korea finished as the third largest consumer of semiconductor materials. Most regions registered high single- or double-digit growth last year.

The Materials Market Data Subscription (MMDS) from SEMI provides annual revenue with 10 years of historical data and a two-year forecast. The annual subscription includes quarterly updates for the materials segment and reports revenue for seven market regions (North America, Europe, Japan, Taiwan, South Korea, China and Rest of World). The report also features detailed historical data for silicon shipments and revenues for photoresist, photoresist ancillaries, process gases and lead frames.

	2021**	2022	YEAR-OVER-YEAR
TAIWAN	$17,715	$20,129	13.6%
CHINA	$12,082	$12,970	7.3%
SOUTH KOREA	$12,134	$12,901	6.33%
REST OF WORLD	$7,896	$8,627	9.3%
JAPAN	$7,275	$7,205	-1.0%
NORTH AMERICA	$5,713	$6,278	9.9%
EUROPE	$3,961	$4,580	15.6%
TOTAL	**$66.776**	**$72,691**	**8.9%**

Source: SEMI (www.semi.org), Juen 2023
Note: Summed subtotals may not equal the total due to rounding.
*Rest of World inculdes Singapore, Malaysia, Philippines, other areas of Southeast Asia and smaller global markets.
**2021 data reflects current updates.

Micron, STEMM Opportunity Alliance and NSF Advance New Workforce Development Collaborations

Micron Technology, Inc., one of the world's largest semiconductor companies and the only U.S.-based manufacturer of memory, announced the Northwest University Semiconductor Network, a partnership focused on collectively developing the next generation of the U.S. semiconductor industry's workforce. The network will drive foundational and emerging research to increase students' opportunities for experiential learning across the semiconductor ecosystem, with equitable access for underrepresented students, including those in rural and tribal communities.

On June 12th, the STEMM Opportunity Alliance (SOA) held its Northwest STEMM Summit: Achieving Equity and Excellence, hosted by Micron, in Boise. At this summit, Erwin Gianchandani, National Science Foundation's assistant director for technology, innovation and partnerships, and Travis York, SOA leader and director of inclusive STEMM ecosystems for equity and diversity at the American Association for the Advancement of Science (AAAS), joined Micron executives as they introduced the network, which includes 13 founding member institutions across six states.

The SOA is a national initiative led by the AAAS, with support from the Doris Duke Foundation, more than a dozen funders and 110+ cross-sector partners that are focused on galvanizing stakeholders to build a STEMM (science, technology, engineering, mathematics and medicine) ecosystem rooted in equity, inclusion and scientific excellence. The SOA was launched at a White House summit in December 2022 alongside advisory council members, including April Arnzen, Micron senior vice president, chief people officer and president of the Micron Foundation. The SOA and Micron

understand that diversity in STEMM is essential to excellence in the scientific ecosystem, which is critical for U.S. economic growth and competitiveness and necessary for building a better society that benefits all communities.

Micron and National Science Foundation (NSF) also announced the next step in a previously disclosed $10 million partnership between the Micron Foundation and NSF. NSF will publish a Dear Colleague Letter with two NSF solicitations that are focused on educator equity and are aligned with the strategic vision laid out in the CHIPS and Science Act. The first of these solicitations, Research Experiences for Teachers (RET), looks to facilitate professional development of K-12 science teachers through research experience. And the second, the Robert Noyce Teacher Scholarship Program (Noyce), addresses the critical need for recruiting, preparing and retaining highly effective elementary and secondary mathematics and science teachers in high-need school districts.

"Partnerships between institutions of higher education, government and industry leaders like Micron are critical so that together, we can build an inclusive and robust talent pipeline equipped with the high-tech skills that will advance U.S. technology innovation and leadership," Arnzen said. "Micron is proud to be expanding our university partnership model through the U.S., leveraging the regional proximity to our leading research and development center in Boise and reaching students that have been traditionally underrepresented in the semiconductor industry, including rural and tribal communities in the Northwest. Our partnership with NSF aims to increase diversity and equity among STEM professionals, helping more students from traditionally underrepresented groups see themselves in their educators, the STEM field and the semiconductor industry. We look forward to furthering the impact of Micron's Northwest University Semiconductor Network and our NSF partnerships as we focus on building a workforce of the future."

Micron plans to invest approximately $15 billion through the end of the decade to construct a leading-edge memory manufacturing fab co-located with the company's R&D epicenter in Boise. This project will create over 17,000 new Idaho jobs, furthering the need for a diverse, highly-skilled workforce. To support the semiconductor industry at scale, Micron has prioritized building relationships with K-12 and higher education partners who can help train students in STEM disciplines, expand research in the broad field of semiconductors and retain a more diverse group of students as they complete their degrees.

"Public-private partnerships help the

Continued from page 6

of 2.4Gb per second, a 50% increase from the previous generation, and approximately 20% increase in read and write speed, the company is confident that it will be able to deliver an improved performance to the smartphone and PC customers using this technology.

Once the product compatibility test with the global smartphone manufacturer is completed, SK hynix will begin supplying the 238-layer NAND product for smartphones, and expand the technology across its product portfolio such as PCIe 5.0* SSDs and high-capacity server SSDs going forward.

"We will continue to overcome NAND technology limitations and increase our competitiveness so that we can achieve a bigger turnaround than anyone else during the upcoming market rebound," said Jumsoo Kim, Head of S238 NAND at SK hynix.

U.S. stay competitive in key technology areas and grow access to high-wage, good-quality jobs in every part of the country," Gianchandani said. "NSF is delighted to partner with Micron to catalyze new career pathways, including through practical experiences for learners of all backgrounds, ages and geographic locations. The new funding opportunity that we are releasing today will broaden participation in the semiconductor workforce and help create opportunities everywhere and unleash innovation anywhere."

"Both the Northwest University Semiconductor Network and the partnership between Micron and the NSF embody the kind of cross-sector and public-private collaboration necessary to achieving equity and excellence in STEMM," York said. "Investing in the critical semiconductor industry and ensuring we have a diverse workforce that can meet the moment is paramount. SOA looks forward to continued work with our diverse group of partners to build a STEMM ecosystem that will maintain American competitiveness and better individual lives."

The founding partners of the Northwest University Semiconductor Network were identified based on their strong collective undergraduate and graduate programs in engineering and other STEM degrees, as well as on their vast R&D expertise and hands-on learning opportunities for students. The network collectively offers access to teaching labs and cleanrooms, a strong foundation of collaboration across institutions, programs connecting to community colleges and strong diversity programs to increase equitable access to engineering education.

The Northwest University Semiconductor Network will expand and prepare the next generation of talent through a framework centered on collaboration, innovation and problem-solving. Micron, in partnership with the network institutions and in alignment with the American Semiconductor Academy and the SEMI Foundation, will champion efforts to modernize and enhance curriculum by sharing best practices and industry-backed technical content, expanding experiential learning programs for greater access to cleanrooms and teaching labs, and bolstering research opportunities for students.

A full list of founding partners of the Northwest University Semiconductor Network follows:

- Boise State University
- Idaho State University
- Montana State University
- Oregon State University
- Portland State University
- San José State University
- University of California, Berkeley
- University of California, Davis
- University of Idaho
- University of Oregon
- University of Utah
- University of Washington
- Washington State University

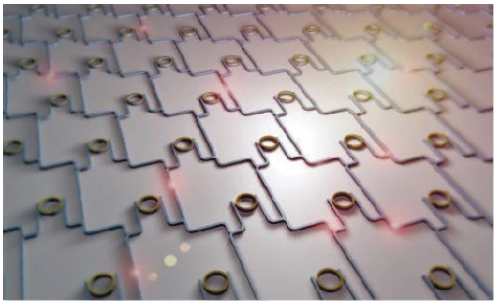
In April 2023, Micron announced its Northeast counterpart network to support company plans to invest up to $100 billion in a megafab in New York.

NEHO: Developing an Artificial Neuron Based on Semiconductor Technology

A tailored combination of material science and photonics to build low energy-consuming neuronal networks is the main goal of the EU-funded project NEHO, coordinated by IIT-Istituto Italiano di Tecnologia (Italian Institute of Technology) in Lecce. The project has been financed with 3 million euros for the next 3 years.

NEHO (Neuromorphic computing Enabled by Heavily doped semiconductor Optics) will exploit semiconductors' properties to create an artificial neuron that can be used to build ultrafast optical neural networks, thus providing a novel and less energy-consuming computational source.

The project will contribute to the



transition from electron-based information processing to an all-photon-based solution: in fact, a photon-based system would, in principle, be much faster, and require a fraction of the energy used by current technologies.

When photons interact with the matter, they generate considerably less heat, but due to their nature, the interaction is weak, such that it is very hard to control the photon flow at small scales. Therefore, NEHO researchers will take advantage from hybrid electron-photon quasi-particles, called *plasmons*, which arises from the interaction of electrons with the light. Plasmons will be produced by using heavily doped semiconductors shone with mid infrared light. Because a plasmon carries both an electron and a photon, one can act on the electronic part, which is generally easier to do, to induce a change on the photonic counterpart. This kind of interaction permits in principle to control the photons at small scales.

The project revolves around the idea of exploiting effects that occur at the surface, rather than the bulk of the semiconductors, as these can be easily modulated by controlling the electron density on the semiconductor surface. Like waves caused by the wind on the surface of the ocean do not require to move water in the deep. This feature will be used to exploit the implementation of a neural network into the development of new machine learning optimization techniques.

NEHO's ultimate goal is to improve the computational power of innovative technologies, such as those based on artificial intelligence algorithms, and at the same time reduce their energy footprint.

"We could revolutionize the way we process information by developing an innovative photonic integrated circuit platform" — explains Dr. Cristian Ciracì, leader of the Computational Nanoplasmonics unit at IIT in Lecce and NEHO's coordinator. "This platform leverages nonlinear photon-plasmon semiconductor technology to allow ultrafast and energy-efficient information processing in the mid-infrared range. With this project, we could be entering a new era of information processing that is faster, more energy-efficient, and more flexible than ever before".

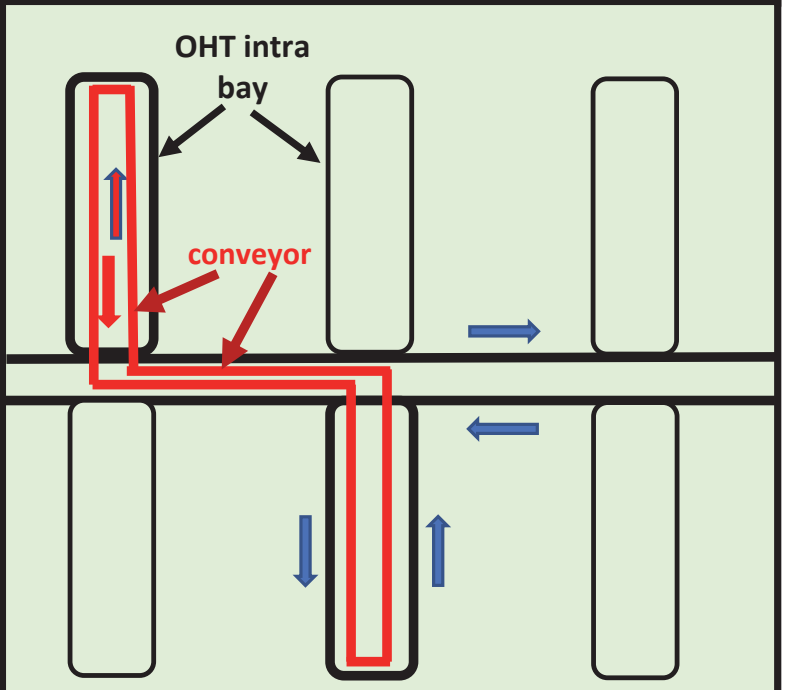
The NEHO European consortium includes Istituto Italiano di Tecnologia (Italy, coordinator), Ludwig-Maximilians-Universität München (Germany), Universiteit Gent (Belgium), CNRS and Université Paris-Saclay (France), Consiglio Nazionale delle Ricerche (Italy).

CONVEYORS FIX AMHS BOTTLENECKS

★FOR BETTER **PROCESS COUPLING**
★REDUCED **CYCLE TIME**
★INCREASED **THROUGHPUT**
★SAVED STOCKER AND EFEM **COSTS**

The Midsx digiflow conveyors seamlessly fit into your existing AMHS Complementing your OHT-s and your MCS with its embedded controls

Call for free solution consulting and cost estimating.



65 YEARS OF MIDDLESEX INDUSTRIES, WRITE or Call:
info@midsx.com or 978-992-2189

Entegris Breaks Ground for Manufacturing Center in Colorado Springs

Entegris, Inc., a supplier of advanced materials and process solutions for the semiconductor and other high-technology industries, held a groundbreaking ceremony in June for its new manufacturing center of excellence in Colorado Springs, Colorado.

At the ceremony, Entegris President and Chief Executive Officer Bertrand Loy noted that the state-of-the-art facility in Colorado Springs will develop products critical to the manufacturing of semiconductors in the United States and around the world. The new leading-edge manufacturing center will serve the semiconductor industry, which is expected to grow to $1 trillion by 2030. Entegris' commitment to innovation and investment in leading-edge manufacturing and supply chain operations will strongly position Entegris — and its customers — for success, and further bolster Colorado's and the U.S.'s position as a technology, manufacturing, and innovation leader across the semiconductor ecosystem.

The new campus, which is targeted to begin initial commercial operation in early 2025, will be built in phases. The 100,000-square-foot facility to be built in the initial phase will support production for Entegris' Microcontamination Control (MC) and Advanced Materials Handling (AMH) divisions. Upon completion, the facility will increase production capabilities for Entegris' most advanced products for filtration and purification as well as wafer carriers, also known as Front-Opening-Unified Pods (FOUPs). The center of excellence will play an important role in the Colorado Springs community and has the potential to create approximately 600 new jobs over several years. Entegris continues to expect to invest approximately $600 million in phases over several years in the facility and is looking forward to working with the city, county, and state to integrate the site and surrounding land into the community in a way that preserves the natural beauty of Colorado Springs.

"The start of construction of our new world-class facility in Colorado Springs represents an important milestone for Entegris," Loy said. "The semiconductor manufacturing industry is set to expand in the United States following the passage of the CHIPS and Science Act, and Entegris' manufacturing center of excellence will enable our organization to play a meaningful role in this important effort. With our large state-of-the-art facility, we will help shorten the supply chain and more efficiently meet the growing needs of domestic manufacturers. We are pleased to have found an innovative and competitive home in Colorado Springs, and we look forward to building on our 30-year history in the region while enhancing our leadership position in the semiconductor ecosystem."

Efabless Reveals Winners of AI-Generated Silicon Design Challenge

Efabless Corporation, the creator platform for chips, announced the winners of its AI Generated Open-Source Silicon Design Challenge. The challenge invited participants to use generative AI to design and tapeout an open-source silicon chip and to do so in three weeks. The wide range of participants and the variety of innovative designs showed how generative AI can accelerate and democratize innovation in chips.

The first-place winner of the contest, by the slimmest of margins, is QTCore-C1 by Hammond Pearce at New York University. The design is a co-processor that can be used for many applications, such as predictable-time I/O state machines for PIO functions as seen on some microcontrollers developed using the Chip-Chat methodology that the NYU team has published.

The second-place winner of the contest is Cyberrio by Xinze Wang, Guohua Yin, and Yifei Zhu at Tsinghua-Berkeley Shenzhen Institute. This design is a RISC-V CPU, implemented with Verilog code produced via a series of prompts given to ChatGPT-4.

The third-place winner of the contest is Asma Mohsin at Rapid Silicon. The design is a Model Predictive Control (MPC) that is used to predict future behavior and optimize control actions for a regulator control circuit provided in MATLAB code to ChatGPT-4 and then implemented with prompts in Verilog.

The winners and all the participants successfully demonstrated how tools such as ChatGPT, Bard and others can revolutionize chip design by automating many of the tedious tasks involved in the development process. Contestants have used these AI tools and others to generate Verilog code from natural language prompts, which they then implemented on the Efabless chipIgnite platform using the OpenLane open-source design flow. The contestants ranged from industry experts to people who had never before designed a chip. None of the winners had ever used the OpenLane Flow before.

Winning designs were selected by a panel of industry-respected judges who used pre-determined criteria, including design completeness, documentation, technical merit, and community interest. Efabless will now fabricate the three winning designs on its chipIgnite shuttle with the winners receiving packaged parts and evaluation boards – a value of $9,750. In addition, the other participants with qualifying designs will receive a free evaluation board and one of the winning AI-generated chips.

"I would like to congratulate the winners for creating such impressive designs in such a short amount of time," said Mike Wishart, CEO of Efabless. "We were struck by the enthusiasm and collaborative participation of the community and the tremendous learning that was achieved by everyone involved. We look forward to our role in facilitating this innovation by providing a forum for learning and a platform to drive AI-assisted development from idea to silicon."

Efabless will soon release information about the second AI-Generated Design Challenge. The challenge will take place over the summer, with tapeouts expected in September.

Quantum-Si Debuts New HQ in Branford, Connecticut

Quantum-Si- Inc. officially debuted its new 29,420 square-foot headquarters in Branford, Connecticut. The new facility extends the company's decade long history in Connecticut, and includes a 7-year lease which will keep Quantum-Si in Branford at least through 2029. The official ribbon cutting, hosted on June 6, featured remarks from Quantum-Si Chief Executive Officer, Jeff Hawkins; Quantum-Si founder Jonathan Rothberg, PhD.; and special guests Governor Ned Lamont, US Senator Richard Blumenthal (D-CT), and Alexandra Daum, the Commissioner of the Connecticut Department of Economic and Community Development (DECD).

The new facility is located at 29 Business Park Drive in Branford, Connecticut, and provides the resources necessary for increased operational scale, development, and collaboration. This milestone represents Quantum-Si's anticipated growth trajectory and ability to support customers and partners using Platinum, the world's first next-generation single-molecule protein sequencing platform.

"The cities of Branford and Guilford, and State of Connecticut have been incredible partners to Quantum-Si since our founding. Our company's roots are here, our relationships are here, and we want to continue building a world-class company here." said Quantum-Si's Chief Executive Officer, Jeff Hawkins. "Our new headquarters reflects the company's ongoing commitment to job opportunities and innovation alongside the local community. We believe New Haven County is the perfect strategic location to attract world-class scientific talent as we seek to grow our teams. We look forward to leading the expansion of the life sciences industry in Connecticut."

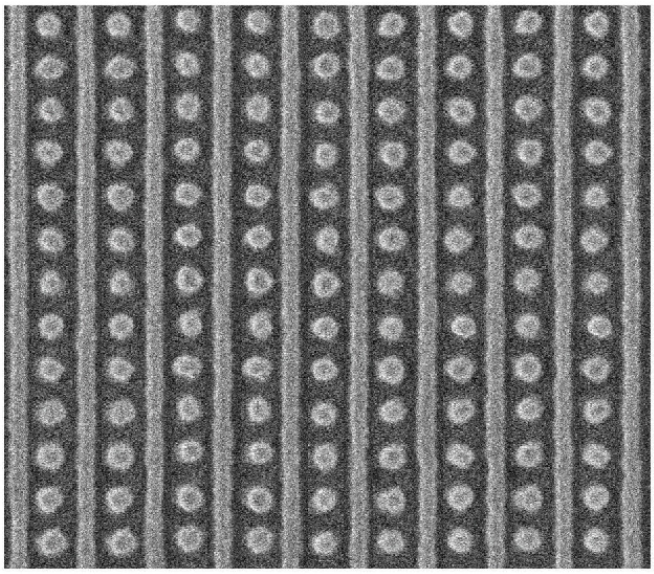
"We have a truly transformative product on the market in Platinum, the next-generation protein sequencer, a leadership bench of highly experienced executives, and now we have the facilities to scale alongside that growth" said Dr. Jonathan Rothberg, Founder and Chairman of Quantum-Si. "Next-generation protein sequencing will be critical for the discovery of a new wave of precision therapies, and I am proud to see this work being done in the community I call home."

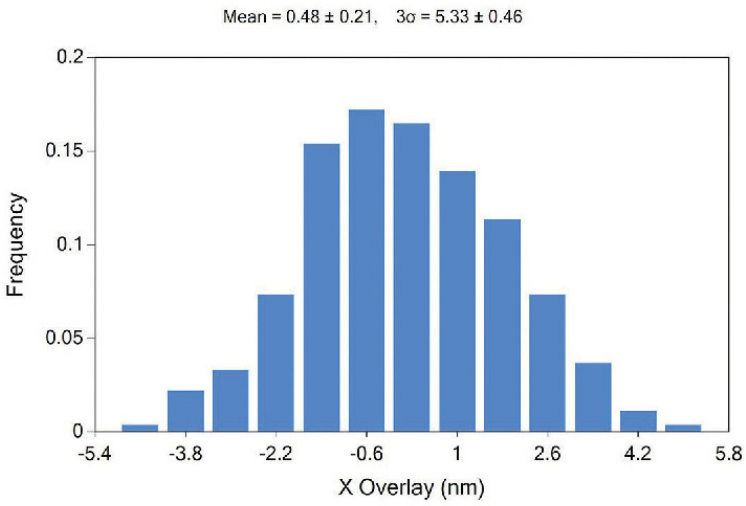
Platinum is the world's first next-generation single-molecule protein sequencing platform that uses proprietary recognizers to identify proteins and amino acids using kinetic binding signatures. Its elegant design and simple workflow enables broad-scale access to proteomic data, for every scientist, everywhere, making it a ground breaking platform for the future of proteomics research. With its small benchtop design and low price-point, Platinum is poised to accelerate breakthroughs across scientific disciplines, from drug discovery to biotech, and help people live healthier and longer lives.

Fractilia Adds Overlay Metrology Capability to Stochastics Control Solutions

Fractilia, the leader in stochastics metrology and control solutions for advanced semiconductor manufacturing, introduced the Fractilia Overlay Package, an optional offering that adds critical new overlay measurement and analysis capabilities to Fractilia's MetroLER™ and FAME™ products. Fractilia's products, which combine the company's patented Fractilia Inverse Linescan Model (FILM™) technology with true computational metrology, are the only proven fab solutions that provide highly accurate and precise measurements of all major stochastic effects – the single largest source of patterning errors at advanced nodes. Fractilia is currently engaged with multiple leading chip manufacturers in analyzing their SEM-based overlay data with the new Fractilia Overlay Package.

In semiconductor manufacturing, overlay is the precise placement of each patterned layer relative to a previous layer to ensure electrical contact and produce a functioning device. Optical-based metrology tools have traditionally been used by chip manufacturers to measure and control this pattern overlay, which is essential to producing high-yielding and high-performing semiconductor devices. These measurements are made on special targets in the scribe line rather than in the device itself. As feature sizes of chips continue to shrink and as stochastic variability has increased with the adoption of EUV patterning, there is a growing offset between what is measured in the scribe line and what is happening in the device. As a result,





The Fractilia Overlay Package measures SEM overlay between features and calculates the overlay distributions, across-wafer variations, and much more. Source: Fractilia.

interest is growing in the use of SEM tools to measure overlay due to their higher resolution and precision compared to optical tools. However, SEMs introduce noise in the SEM image, which can be easily confused with stochastic variability on the wafer.

"SEM-based overlay measurements are increasingly being used to improve advanced patterning process control, but they are influenced by the same random and systematic SEM errors that affect other stochastics measurements," stated Fractilia CTO Chris Mack. "Fractilia is unparalleled in measuring and subtracting out SEM noise through our proven FILM technology, and as a result, our customers are turning to Fractilia to help them apply our technology to improve the accuracy of their SEM overlay measurements. In addition, by combining SEM stochastics measurements with optical overlay measurements through our Fractilia Overlay Package, we believe that we not only can improve SEM overlay metrology accuracy, but also provide better lot dispositioning and correctables, which in turn can improve patterning control and reduce non-zero offset (NZO) or its variability."

"Unbiased" measurements provide more accurate description of what's on the wafer

Fractilia's FAME portfolio of solutions uses a proprietary and unique physics-based SEM modeling and data analysis approach that measures and subtracts the random and systematic errors from SEM images to provide measurements of what is on the wafer rather than what is on the images. FAME measures all major stochastic effects simultaneously, including line-edge roughness (LER) / linewidth roughness (LWR), local CD uniformity (LCDU), local edge placement error (LEPE), and stochastic defects, as well as provides CD measurements. It provides the highest signal-to-noise edge detection in the industry (up to 5x higher signal-to-noise ratio than other solutions), and extracts more than 30x more data from each SEM image.

With the new Fractilia Overlay Package, Fractilia adds highly accurate SEM-based overlay measurements, including their stochastic properties, on top of its existing measurement capabilities. Fractilia's products have been adopted throughout the industry by dozens of companies, including leading semiconductor manufacturers, equipment companies, materials suppliers and research organizations.

Novel Ferroelectrics for More Efficient Microelectronics

As computational energy usage continues to grow, it is on pace to potentially become the leading source of energy consumption in this century. Memory and logic are physically separated in most modern computers, and therefore the interaction between these two components is very energy intensive in accessing, manipulating, and re-storing data. A team of researchers from Carnegie Mellon University and Penn State University is exploring materials that could possibly lead to the integration of the memory directly on top of the transistor. By changing the architecture of the microcircuit, processors could be much more efficient and consume less energy. In addition to creating proximity between these components, the nonvolatile materials studied have the potential to eliminate the need for computer memory systems to be refreshed regularly.

Their recent work published in *Science* explores materials that are ferroelectric, or have a spontaneous electric polarization that can be reversed by the application of an external electric field. Recently discovered wurtzite ferroelectrics, which are mainly composed of materials that are already incorporated in semiconductor technology for integrated circuits, allow for the integration of new power-efficient devices for applications such as non-volatile memory, electro-optics, and energy harvesting. One of the biggest challenges of wurtzite ferroelectrics is that the gap between the electric fields required for operation and the breakdown field is very small.

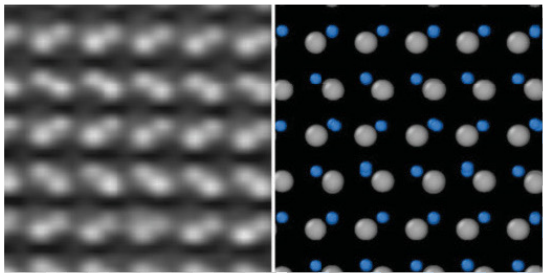
"Significant efforts are devoted to increasing this margin, which demands a thorough understanding of the effect of films' composition, structure, and architecture on the polarization switching ability at practical electric fields," said Carnegie Mellon post-doctoral researcher Sebastian Calderon, who is the lead author of the paper.

The two institutions were brought together to collaborate on this study through the Center for 3D Ferroelectric Microelectronics (3DFeM), which is an Energy Frontier Research Center (EFRC) program led by Penn State University through funding from the U.S. Department of Energy's (DOE) office of Basic Energy Science (BES).

Carnegie Mellon's materials science and engineering department, led by Professor Elizabeth Dickey, was tapped for this project because of its background in studying the role of the structure of materials in the functional properties at very small scales through electron microscopy.

"Professor Dickey's group brings a particular topical expertise in measuring the structure of these materials at very small length scales, as well as a focus on the particular electronic materials of interest of this project," said Jon-Paul Maria, professor of Materials Science and Engineering at Penn State University.

Together, the research team designed an experiment combining the strong expertise of both institutions on the synthesis, characterization and theoretical modeling of wurtzite ferroelectrics. By observing and quantifying real-time polarization switching using scanning transmission electron microscopy (STEM), the study resulted in a fundamental understanding of how such novel ferroelectric materials switch at the atomic level. As research in this area progresses, the goal is to scale the materials to a size in which they can be used in modern microelectronics.



In-situ experimental STEM images (left panel) and first-principles calculation prediction (right panel) France.

HYBRID AMHS
FOUP/SMIF OHT

For more information, visit our website by scanning the QR Code or email us at sales@stratusauto.com

STRATUS Automation

www.stratusauto.com

Copyright © 2023. Stratus Automation Sdn Bhd. All rights reserved.

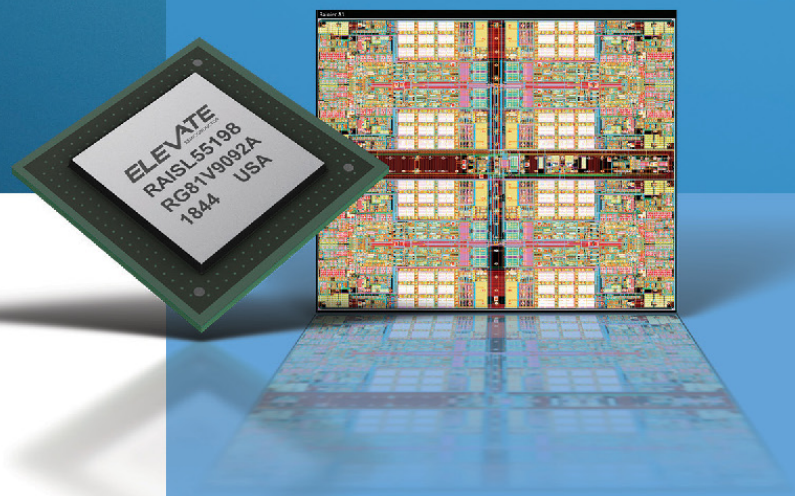
Visit us at:

SEMICON® WEST

July 11-13, Booth 1835

ATE is *Our* Business

ElevATE: High-Density, Low-Power Semiconductor Test Solutions



ELEVATE SEMICONDUCTOR

www.elevatesemi.com

As technology advances at an unprecedented pace, the demand for semiconductor test solutions surges. Elevate is dedicated to addressing the challenges of power efficiency, density, speed, and voltage. Are you prepared to put your upcoming product roadmap to the test?

Challenges in Testing Semiconductors with 5nm and Smaller Process Geometries

ELEVATE SEMICONDUCTOR

Elevate Semiconductor's latest DPS/PMU products offer improved architectures to maximize power efficiency, while providing high channel counts and DC accuracy necessary to test tomorrow's state-of-the-art devices.

OVER RECENT YEARS, THE SEMICONDUCTOR market has seen considerable growth, driven largely by the advancements in AI, 5G, HPC, Edge Computing, and autonomous vehicles sectors. This has led to a surge in demand for new fabs capable of supporting increasingly small process geometries.

Current and projected fab investments indicate a trend towards sub-5nm and in some instances, sub-2nm process nodes. Driven by the CHIPS Act and other non-US funding, more than $150 billion worth of fabs are presently under development, with industry leaders such as Samsung and TSMC opening fabs in the US, and Intel also investing in European fabs.

Next-generation products, designed with these smaller geometries, are also under development. Examples include the next-generation Apple iPhone 15 Pro/Ultra 3nm processor and Nvidia's H100 GPU/AI 4nm processor.

However, these advancements bring new complexities. Testing semiconductors with process nodes of 5nm or smaller is a challenging task. Manufacturers are in pursuit of future-proof test solutions that ensure optimal yield and quality while also controlling costs.

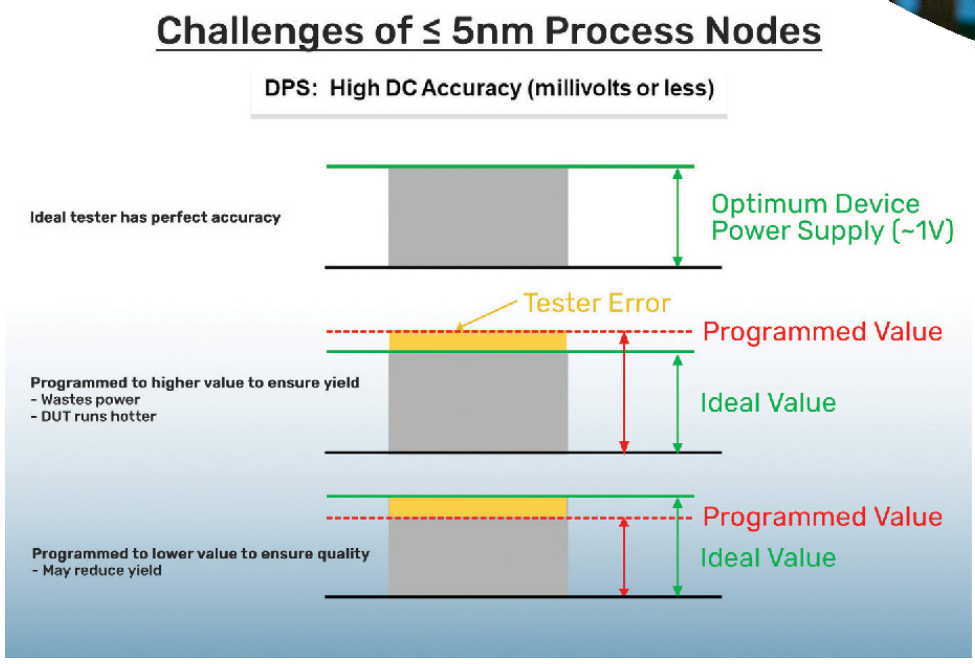


Figure 1. High DC Accuracy of 1 millivolt or less is often required.

However, as semiconductors become smaller with more gates and more complexity, testing takes longer. Without improvements in testing, these costs can become the largest part of a company's cost of goods.

Testing challenges of sub-5nm process nodes

Components designed with sub-5nm process nodes demand reduced operating voltages (less than 1 volt) while drawing 1000s of amps (**FIGURE 1**). These tiny geometries operate at faster speeds and are densely packed with features and functionalities. The chips can be large in format and have pin counts exceeding 1000, all of which need to be tested before final production. This can cause considerable issues. In the case of large AI processors, it's impossible to conduct a

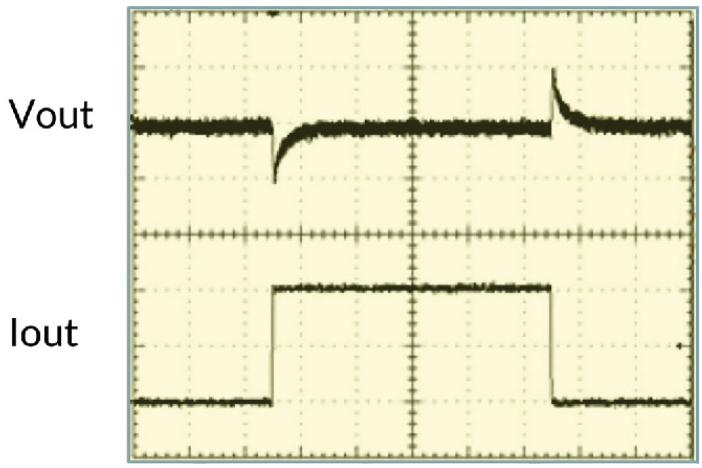


Figure 2. Voltage Droop caused by large current and slow load transient response.

full test all at once with current testing technology. The only feasible option is to partially test them in series, one section at a time.

In order to test these sophisticated chips with so many pins, the test solutions need to have numerous channels, positioned as closely as possible without causing overheating. Enhanced power efficiency is not only necessary for power savings but also to enable increased density. With operating voltages at 1 volt or less, high DC accuracy of 1 millivolt or less is essential to provide the proper voltage to the device under test (DUT).

Other requirements include a faster load transient response to avoid potential yield issues caused by voltage droop (**FIGURE 2**).

What is a DPS/PMU?

A DUT Power Supply (DPS) is a semiconductor that provides the supply voltage required for the Device under Test (DUT). The DPS can be operated in both force voltage (FV) and force current (FI) mode and has functionality similar to a typical power supply as well as the capability to read back and measure the voltage and current. A Parametric Measurement Unit (PMU) and a DPS are alike, but the DPS has higher current capabilities than the PMU. A Per-pin PMU (PPMU) architecture significantly reduces test time as every DUT pin can be tested simultaneously.

DPS/PMU Solutions for Testing ≤5nm process nodes from Elevate Semiconductor

Elevate Semiconductor, established in 2012, is dedicated to developing cutting-edge technology that facilitates the next generation of testing in the semiconductor industry. The company offers a comprehensive range of Device Power Supply (DPS) and Parametric Measurement Unit (PMU) products, designed to last and keep up with technological advancements (**FIGURE 3**).

Elevate Semiconductor's latest DPS/PMU products offer improved architectures to maximize power efficiency, while providing high channel counts and DC accuracy necessary to test tomorrow's state-of-the-art devices.

Unique in the industry, Elevate Semiconductor focuses solely on testing. Their test solutions include a full range of DPS/PMU, Pin Electronics, and VI products. The company's solutions are designed to be low-power and high-density, aiming to reduce testing costs (through parallelization).

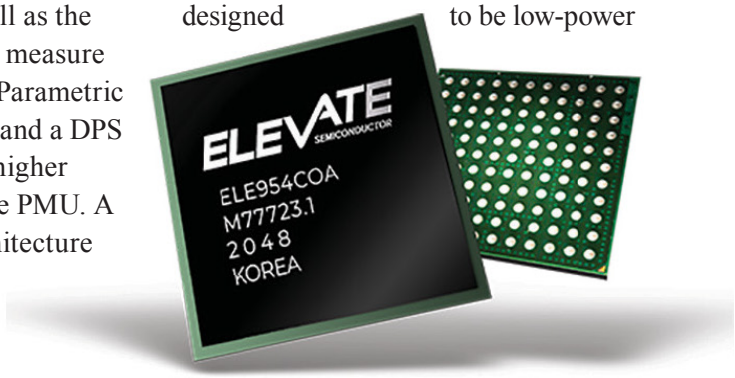


Figure 3. Elevate next-gen D.PS/PMU.

To learn more about Elevate Semiconductor and their DPS/PMU solutions, go to https://elevatesemi.com/engage/

Tim Bakken, Elevate Semiconductor Vice President of Global Sales stated: "By actively engaging and collaborating with our clients, Elevate fosters a deep understanding of the problems our customers are trying to solve. This collaborative approach drives our ability to develop effective solutions for every challenge. Our extensive portfolio includes a wide range of offerings, from standard solutions to semi-custom and full-custom options. With our customer-centric focus, collaborative mindset, and diverse offering, Elevate Semiconductor is committed to delivering exceptional products and service to our valued customers.

The State of the Semiconductor Industry 2023

PETE SINGER, Editor-in-Chief

New approaches will be required to address challenges associated with energy efficiency, the explosion of AI and emerging automotive applications.

AMID A BACKDROP OF SEVERAL SEVERE global crises — war, climate change, an aging population and supply chain disruption, plus the need for better mobility, reliable energy, healthcare — the role that semiconductors play on the world stage has never been more important. "It is clear that complex deep tech innovations will be essential to address these phenomenal challenges our planet and our society is facing," said Luc Van den hove, president and CEO of imec, speaking at the consortiums recent ITF World meeting, held May 17-18, in Antwerp, Belgium. Other presenters included Jensen Huang, president, CEO and Member of the Board of Directors at NVIDIA, Mark Papermaster, executive vice president and chief technology officer of technology and engineering at AMD, Jaihyuk Song, corporate president and CTO of Samsung Electronics' Device Solutions Business, and Christophe Fouquet, CBO & Board of Management at ASML. In this article, we summarize their insights into the major trends, challenges and possible solutions in the years ahead for the semiconductor industry.

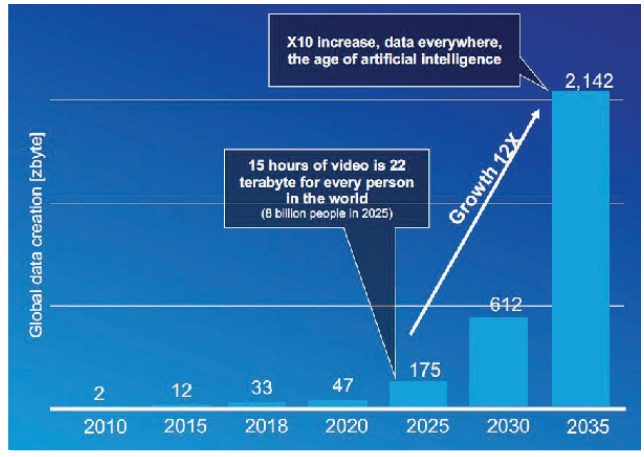


Figure 2. Data creation is expected to increase 12X in the next decade. *Source: ASML, ITF World.*

More data than stars

There are many drivers of semiconductor consumption — cloud computing, 5G, IoT, automotive, mobile, AR/VR, phones, etc. — and, in just the last year, generative artificial intelligence and ChatGBT. The general consensus is that the market for semiconductors will reach $1Trillion by 2030, practically double was it is today. To achieve this, billions of dollars in investments in new factories and research centers have been announced over the last year, spurred on by billions in government incentives in the U.S. (The CHIPS Act), Europe, Japan, Korea and China. Major efforts are underway to supply the workforce that will be needed.

One of the most useful ways of looking at semiconductor demand is the amount of data being generated, which of course, must be collected, analyzed and acted upon. "The amount of data generated each year is exploding," said Samsung's Song. "It is more than two and a half times the number of stars in

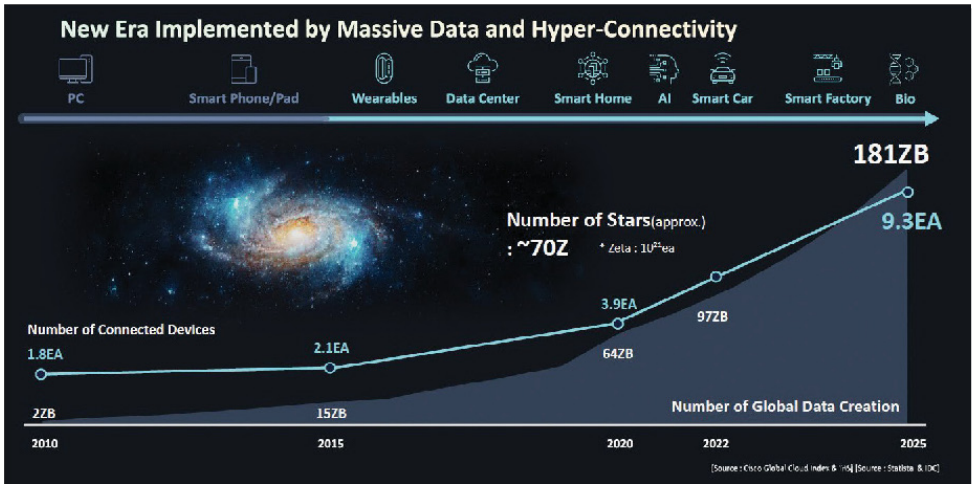


Figure 1. The world's data creation has already exceed the estimated number of stars in the universe. *Source: Samsung, ITF World.*

semi™

Connecting Semiconductors and Electronics

About SEMI:
SEMI connects more than 2,500 member companies and 1.3 million professionals worldwide to advance the technology and business of electronics design and manufacturing. The breadth and depth of our events, programs and services help members accelerate innovation and collaboration on the toughest challenges to growth.

ESDA Electronic System Design Alliance

FlexTech

FOA Fab Owners Alliance

MEMS & SENSORS INDUSTRY GROUP℠

SOI Industry Consortium

the entire universe currently estimated by space scientists."

This trend is illustrated in **FIGURES 1** and **2.** AMD's Papermaster noted that if we stay on the same pace, the power needed for supercomputer installation will approach that of a nuclear power plant (**FIGURE 3**). "That's a pretty scary data point, and what it really tells you is that we are hitting absolute limits of energy consumption," he said. The solution: changing the way computing is done to bring the memory closer to the CPU, and new more energy efficient chip designs and architectures.

"It will take a holistic approach to bend the curve, to be able to improve our rate of energy efficiency," Papermaster said. "We have to expand the sphere of co-operative and more co-design optimizations across the stack.. and really drive system level implementations. We have to think differently, and really leverage innovation. If you look at the way that we've always designed — focusing on silos and then putting that sum of parts together — it simply can't bend the curve," he said.

AI is set to worsen the problem. "AI, with the large language models and the energy needed to train, is going to actually grow that rate and pace of energy consumption," Papermaster said, noting billions of parameters are used to today to train models such as ChatGBT3, and "it's heading to trillions of parameters."

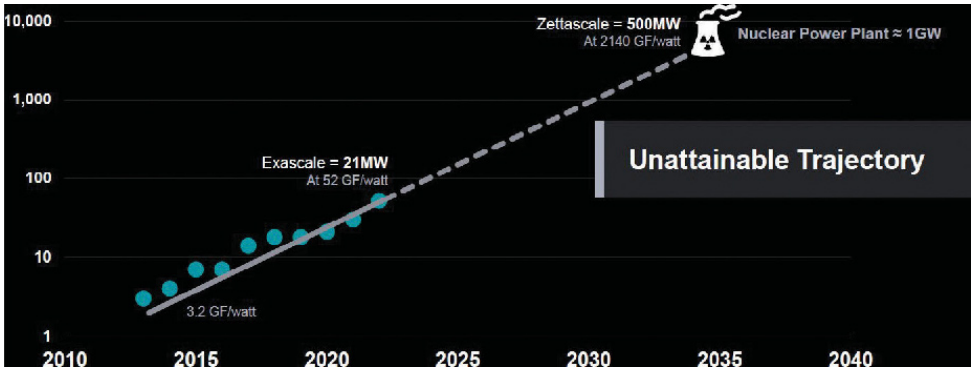


Figure 3. Green500 supercomputer GFLOPS/Watt and projection. The Green500 is a biannual ranking of supercomputers in terms of energy efficiency. *Source: AMD, ITF World.*

The explosion of AI

Van den hove said the fast growth of generative AI is driven by three primary factors: Advancement in AI models, access to large amounts of data, and massively increased compute power. "The compute needs for AI are literally exploding, and we are just at the beginning," he said. "In order to generate this required compute power in a sustainable way, we will need drastically more performant devices and systems."

Nvidia's Huang noted that, in the past decade, the scale of deep learning has increased by a million times. "Then, ChatGPT arrived.. the AI heard around the world," he said. "ChatGPT, a generative pre-trained transformer, is a large language model with hundreds of billions of parameters trained on trillions of words and sentences. It has learned the representation of human language and can generate text. ChatGPT is revolutionary due to its ease of use and incredible capabilities. It's a glimpse into the future of computers that understand any programming language and perform a broad range of tasks."

Huang noted that, already, over a thousand generative AI startups are inventing new applications. Tabnine, for example, is a contextual code assistant that helps complete lines of code or generate entire functions from a prompt description. It can program in multiple popular languages from JavaScript and Python to Rust, GO, and BASH.

Runway uses generative AI to create and edit images and videos. Its production quality is so impressive that several Oscar nominated films have used it. With Runway's, generative AI, anyone can tell stories through pictures and video.

Insilico Medicine uses generative AI to create a candidate drug in one third a time, and one tenth the cost of traditional methods, which usually take over four years and cost around $500 million. Insilico Medicine is used by over 20 pharmaceutical companies.

"The market impact of perception AI is already significant," Huang said. "The impact of generative AI will be much bigger. The remarkable ease of use allowed ChatGPT to reach over 100 million users in just a few months, making it the fastest growing application in history." He said that every form of information will be understood and enhanced by generative AI, "from human language, music, pictures, video and 3D to genes, proteins, and chemicals." The content on 200 million websites will be personalized and generated by AI, he said, and billions of customer service calls will be automated by AI. Generative AI will also assist 25 million software developers and hundreds of millions of creators.

What is the next wave of AI? Huang said it's "Embodied AI" which refers to intelligent systems that can understand, reason about, and interact with the physical world. Examples include robotics, autonomous vehicles, and even chatbots that are smarter because they understand the physical world.

Automotive requirements

Automotive applications are another hot driver of semiconductor technology, including electric vehicles (EVs) and autonomous driving. In addition to advanced battery management systems, the future car requires next generation sensing,

45th EOS/ESD SYMPOSIUM AND EXHIBITS

October 1-6, 2023
Riverside Convention Center
Riverside, CA

Program and registration details updating at:
https://esda.events/

Our program is packed! While we prepare for the full program, check out the incredible tutorials available this year and the deep discounts for registering early for the symposium!

Tutorials offered:

FC340: ESD Program Development and Assessment – ANSI/ESD S20.20 Seminar-Two days ($1,710)

FC100: ESD Basics for the Program Manager ($950)

FC101: How To's of In-Plant ESD Auditing and Evaluation Measurements ($950)

DD/FC240: System Level ESD/EMI (Principles, Design Troubleshooting, & Demonstrations) - ($585)

DD134: Fundamentals of ESD System Level ($585)

Symposium pricing:

ESDA Member: $600
Changes to $800.00 after Monday, July 24, 2023

ESDA Life Member Fee $525.00
Changes to $725.00 after Monday, July 24, 2023

Non-ESDA Member Fees $700.00
Changes to $800.00 after Monday, July 24, 2023

Membership discounts apply to current members as of May 31, 2023. Memberships processed after this date will not apply.

Register at:
https://cvent.me/G57zBw

combining multiple multispectral cameras and multiple radars, including short range and long range radars. "All these sensors will generate incredible amounts of data," Van den hove said. "To process this, our cars will need radically new compute architectures implementing massive amounts of AI running on powerful, centralized electronic control units. When evolving from current ADA systems towards increasing levels of autonomous driving, the required compute power is increasing exponentially. The electronic control units in the cars are actually quickly growing into what used to be a supercomputer performance level," he said.

The energy efficiency of these compute platforms will also become extremely important. "Power consumption will explode and could contribute to consuming up to even one third of the battery range when navigating in busy city centers," Van den hove said. "The chips needed for these ECUs will be among the most complex monolithic designs one can find and will be built using the most advanced silicon nodes."

He added that, while the semiconductor content in cars will increase tremendously leading to a growing semiconductor demand, the total number of globally produced cars is not increasing significantly. "Designing these ultra expensive compute systems on chip and optimizing them in a flexible way for specific markets like high-end and low-end markets is really not obvious," he said. "We will therefore need to optimize these compute architectures taking into account flexibility, complexity, energy efficiency, and affordability."

One approach that's gaining interest in this regard is chiplet-based designs, where the design is broken out into smaller pieces of silicon chiplets that are fused together using advanced heterogeneous integration techniques. "Using chiplets, ECUs can be configured in a more flexible way, allowing for high-end or low-end configurations without having to redesign the entire chip. It'll allow for the reuse of IP as the chips will be smaller, it'll provide better yield and reliability, which is essential for the auto industry," Van den hove said.

Moore's Law, 2D scaling, 3D stacking and chiplets

The semiconductor industry has enjoyed the benefits of 2D scaling, as defined by Moore's Law, for decades. As that type of scaling slows, alternative approaches, such as chip stacking and chiplets are coming into play. "We believe that Moore's law, as we traditionally know it, will still be essential. Indeed, feature-based scaling by reducing the dimension of our individual transistors is still one of the only techniques that provides true exponential complexity increase," Van den hove said.

He showed imec's roadmap for the next decade (**FIGURE 4**), which includes the continued use of 2D scaling, as well as new technologies, such as forksheet transistors and complementary FET (CFET). In a CFET architecture, n- and pMOS devices are stacked on top of each other, allowing further maximization of the effective channel width. "We can enable continued scaling for many more generations by combining lithography-based scaling and the introduction of new interconnect schemes, new materials, and new device architectures, such as the CFET device," Van den hove said.

He also spoke of "walls" that may be difficult to get over. The power wall, where it becomes increasingly difficult to scale the power consumption per transistor, and the memory wall, defined by the delay between the time the CPU needs data from SRAM-based caches or DRAM, and the moment it can access it. "With a lot of AI workloads, it is getting harder to get data efficiently in and out of our processors. In fact, for specific AI workloads in high performance computers, in some cases processors are only active less than 3% of the time," he said. "Most of the time they're actually waiting for data, so we'll, therefore, need to reconsider memory system architecture."

Imec has also identified three other walls:

- **The scaling wall**: pure lithography-enabled scaling is slowing down. It is becoming increasingly difficult since individual structures of microchips and transistors are approaching the size of atoms, where

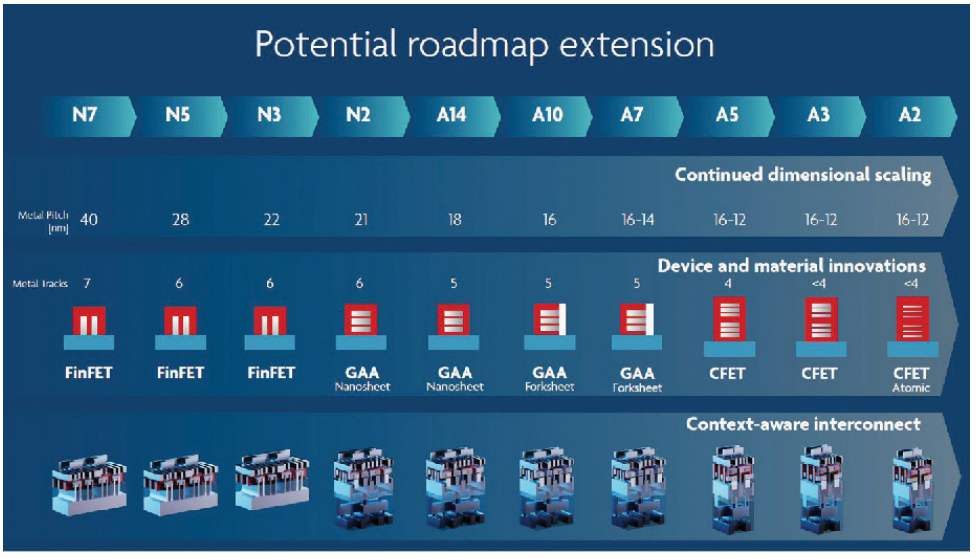


Figure 4. Imec's new roadmap includes the continued use of 2D scaling, as well as new technologies, such as forksheet transistors and complementary FET (CFET). *Source: imec, ITF World.*

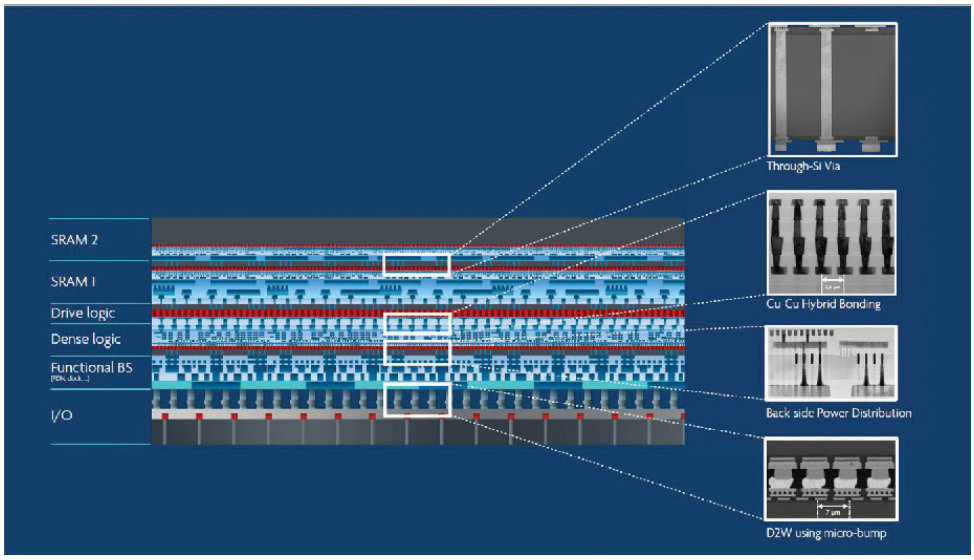


Figure 5. Advanced wafer stacking techniques could employ through-silicon via technology and copper-to-copper hybrid bonding to enable very high dense interconnects between the logic and the memory cache layers. *Source: imec, ITF World.*

quantum effects begin interfering with the operation of microchips.

- **The sustainability wall**: the manufacturing of semiconductor devices contributes to an increasing environmental footprint, including greenhouse gas and water, natural resources, and electricity consumption.
- **The cost wall**: obviously, chip manufacturing costs may explode with the complexity increase, together with the costs for design and process development.

Van den hove said the industry will need a wide portfolio of technology options that will allow for specific technology combinations for addressing specific workloads and system requirements. "Interposer chips and tiny silicon bridge interposers will allow for very fast connections between multiple core processor chiplets and 3D DRAM memories," he said.

Today, DRAM memories are typically built up as a 3D stack of layers, while the core processor chiplets are made using a single die, a monolithic 2D system of chip. "The next way to enable heterogeneous integration is to build up the advanced 2D processor die as a 3D stack of layers, splitting up the functionality across multiple layers. This can be done using advanced wafer stacking techniques such as by using through silicon via technology, copper-to-copper hybrid bonding, enabling very high dense interconnects between the logic and the memory cache layers," Van den hove said. **FIGURE 5** illustrates how this might look in practice.

"We believe that this will be one of the ways to extend Moore's law much further than what we believe possible. We refer to this roadmap as CMOS2.0, realizing the next phase of Moore's law by using 3D stacking to enable further 2D scaling. 3D alone will bring us only linear complexity increase," said Van den hove.

Another issue is that the power consumption of advanced chips is increasing to a point where there are some limits on extracting the heat out of the chip. "It is clear that in order to enable such complex 2D, 3D stacked core processor architecture, we will need to address the heating problem and learn how to cool the chips more effectively. For this, we're developing innovative cooling systems, which can be fabricated using 3D printing techniques," Van den hove said. This direct liquid jet impingement cooling results in much higher heat extraction efficiencies, up to five times better than the current state-of-the-art.

Samsung's Song identified trends specifically related to logic, DRAM and NAND device structures. In logic, to resolve the problem of cell height scaling, he said Samsung is developing backside interconnects. "In addition, innovation in structure and materials such as channel materials with optimum layer thickness can have excellent gate controllability compared to silicon channels," he said (see **FIGURE 6**).

In DRAM, to address the issue of lateral scaling, Samsung is studying a vertical channel stacked 3-D structure that Song said is very different from the existing 2-D structure. "We are also trying to introduce all low-leakage channel material, not the silicon channel that has been used so far leakage current control," he said. "In order to solve the cell capacitor problem, we are making approach to overcome the limitations by studying a new concept of DRAM without a capacitor."

In NAND, technologies such as multi-stack will continue to be developed. "New materials would be required to overcome mechanical stress or sensing margin. And new concepts of device development that can lower operating voltage would be required to further shrink up the unit device size," he said. "To increase the resolution and uniformity of EUV, the high-NA EUV technology, which uses larger mirrors is also needed."

The need for innovation and collaboration

"In order to continue to develop the digital transformation era, the technology for storing and processing such big data requires a new era of development, and there is a constant need for innovation in the semiconductor industry to make this possible," Song said. Continued on page 35

Carbon Capture, Utilization, and Storage Technologies' Role in Fighting Global Warming

CHRIS JONES, Edwards Vacuum

Carbon capture, utilization, and storage (or sequestration) (CCUS) refers to various technologies that capture carbon dioxide at a source or from the atmosphere (direct air capture) and then either use it or store it so that it does not contribute to warming the planet.

THERE IS BROAD CONSENSUS WITHIN THE scientific community that human-generated greenhouse gas (GHG) emissions are causing global warming. If we can emit them, might we capture and remove them? The answer is yes, though perhaps not as easily. But capture them we can, and we are. The real questions are, can we do it fast enough and big enough to make a difference, and can we afford it, in terms of dollars and energy? These questions fall within the purview of geoengineering, the third leg (in addition to mitigation and adaptation) of the framework within which most experts discuss as a response to climate change. Carbon capture, utilization, and storage (or sequestration) (CCUS) refers to various technologies that capture carbon dioxide at a source or from the atmosphere (direct air capture) and then either use it or store it so that it does not contribute to warming the planet.

We are currently emitting about 55 gigatons of GHG (CO_2 equivalent – CO_2e) per year. CO_2 is the major component, but there are also major contributions from other gases (**FIGURE 1**). Methane (CH_4) and nitrous oxide (N_2O) emissions are smaller in volume but have much greater global warming potentials, ton for ton, over a 100-year period (GWP-100), 27.9X and 273X the GWP-100 of CO_2 (respectively). Fluorinated gases also contribute significantly to warming with GWP-100 as much as 25,000X CO_2. At the current rate of emission, we will exceed the budget required to restrict warming to less than 1.5°C before 2030.

In the 2015 Paris Agreement most countries in the world agreed to reduce GHG emissions enough to hold the increase in the global average temperature to well below 2°C above pre-industrial levels and pursue efforts to limit the

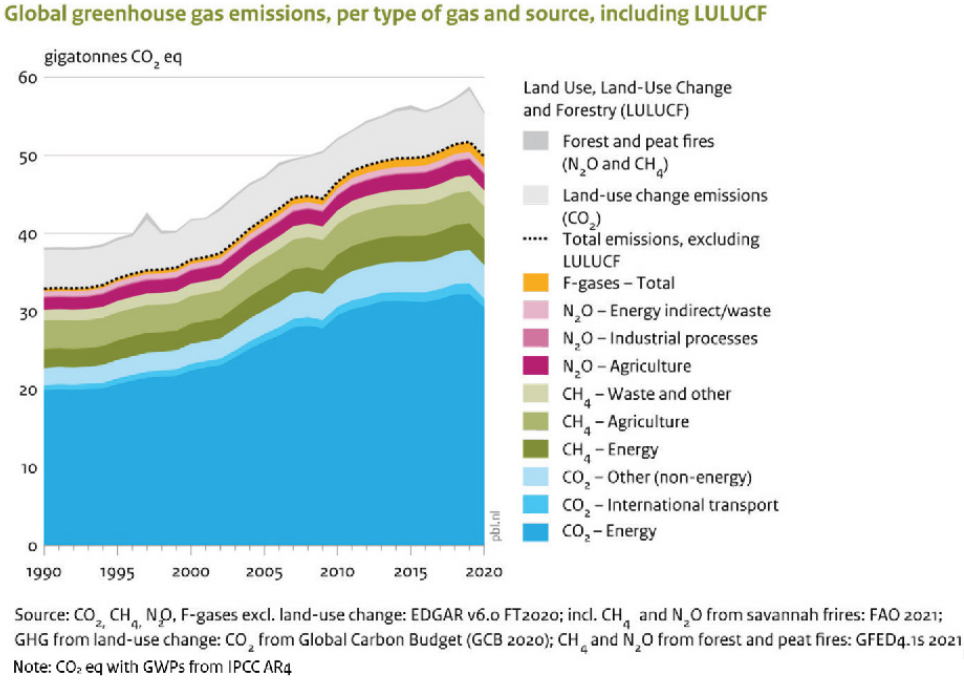


Figure 1. Global GHG emissions by type of gas and source. CO2 from energy production is by far the largest contributor. (https://www.pbl.nl/en/publications/trends-in-global-co2-and-total-greenhouse-gas-emissions-2021-summary-report)

temperature increase to 1.5°C above pre-industrial levels. In 2018 the Intergovernmental Panel on Climate Change (IPCC) warned that we would see substantial impacts if the temperature increased by more than 1.5°C. The Science Based Targets Initiative (SBTi) establishes reduction protocols and standards for companies to use to measure GHG emissions against targets that will limit warming to 1.5°C (**FIGURE 2**). To do so we must approximately halve emissions each decade, i.e., 50% by 2030, then another 50% by 2040, before achieving net-zero emissions by 2050. Emissions from some sectors, such as air transport, may be difficult or impossible to eliminate. CCUS provides a means to remove residual CO_2 emissions that cannot be completely eliminated, especially as we come closer to the end of the journey where progress from reductions becomes harder to realize.

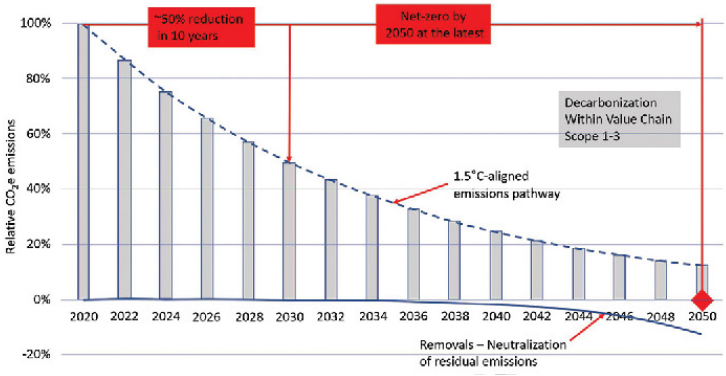


Figure 2. Reductions in GHG emissions needed to limit warming to 1.5°C.

There has been accelerating growth in CCUS efforts around the world over the last few years. Early results from some innovative technologies are encouraging, and it may be that CCUS will play a larger role in the fight against warming than originally suggested. **FIGURE 3** summarizes CCUS projects in development or operation worldwide as of September 2021. Both the number of projects and their geographical distribution are increasing. Net maximum capacity (not actual operating capacity) was 175 MMT per year. That is a respectable number, but it is less than 0.4% of global emissions.

Capture

Carbon capture can contribute in several ways to our efforts to achieve net-zero emissions. Most immediately, it can reduce emissions from fossil-fuel based electrical power generation that accounts for more than 60% of emissions in the U.S. We are making progress in developing green energy sources, including solar, wind, nuclear and geothermal, but capturing carbon from fossil fuel combustion would allow a slower phase out of fossil fuels while still meeting emission targets. In so doing it would ease the transitional impact on employment in the energy industry and avoid the cost of retiring fossil fuel facilities and infrastructure prematurely.

In other industries, carbon emissions come from the process itself. Examples include the production of cement and the manufacture of iron and steel. In the first step of the cement process, limestone (primarily calcium carbonate) is heated to form calcium oxide (lime) and CO_2. Although new cement formulations can reduce process emissions, capturing the carbon from this step is the only practical way to reach net-zero in this sector. Likewise, the initial step in the manufacture of steel involves the reduction of iron in iron ore (primarily iron oxide) by exposing it to heat and carbon monoxide, producing metallic iron and CO_2. There are other ways to reduce iron, such as using hydrogen instead of carbon monoxide as the reducing agent, but they are not yet widely used. Both cement and iron/steel making require high temperatures that are usually supplied by fossil fuel combustion.

The third option is capturing and storing carbon directly from the atmosphere. The primary challenge for direct air capture is the very low concentration of CO_2 in the atmosphere, which makes the capture process energy intensive. Unless the energy it uses is renewable, direct air capture can add more CO_2 than it removes. Still, it is the only option, of the three discussed here, that can actually take CO_2 out of the atmosphere.

Bioenergy with carbon capture and

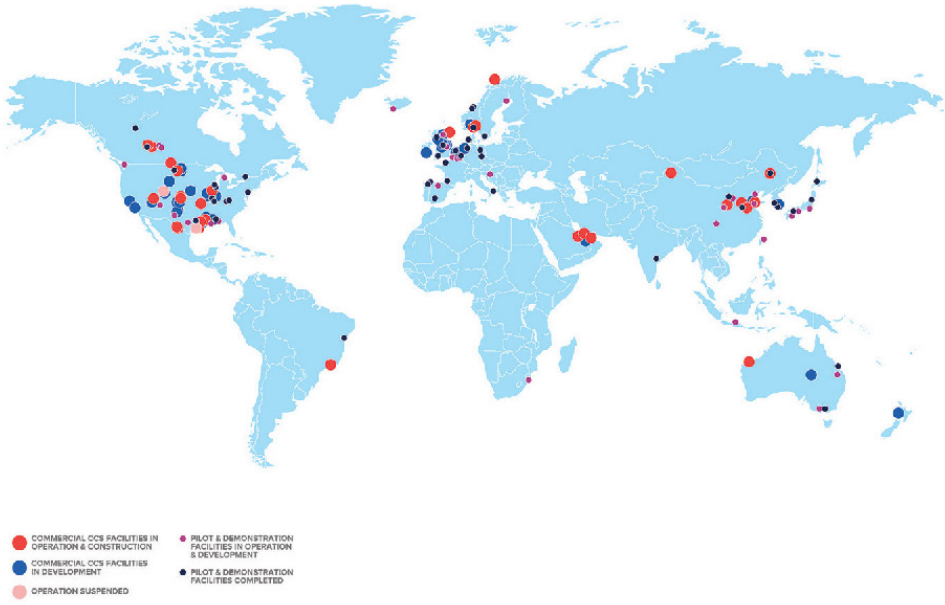


Figure 3. Relative size, status, and location of commercial CCUS facilities

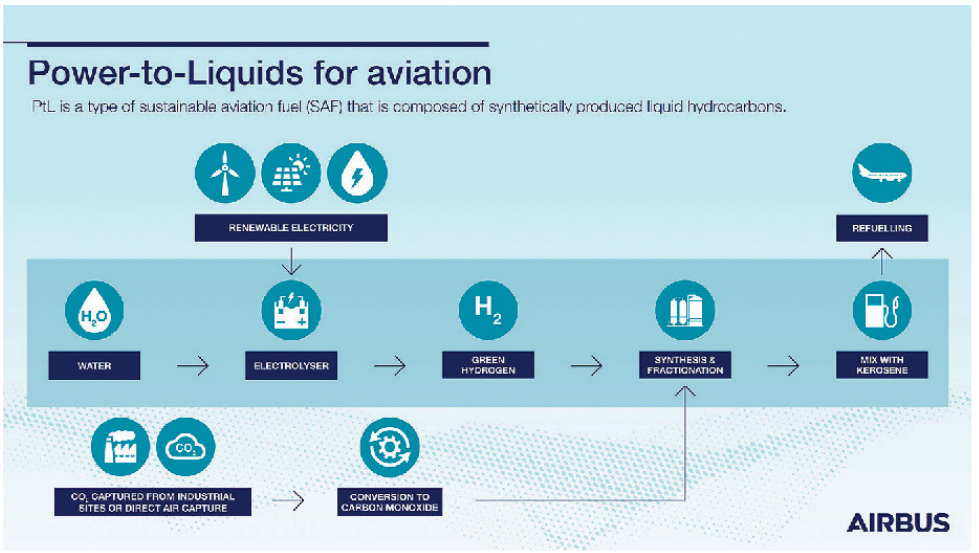


Figure 4. Energy pathway for producing sustainable aviation fuel

storage (BECCS) refers to energy pathways that capture and store biogenic carbon. For example, in the production of ethanol from corn, the growing corn captures CO_2 directly from the atmosphere and converts it to carbohydrates. A fermentation process converts the carbohydrates to ethanol and CO_2. The ethanol may be added to fuels and returned to the atmosphere when it is burned, or it may serve as feedstock for other industrial processes. The CO_2 generated by fermentation is highly concentrated and, therefore, easily captured and stored in geological formations. Other, more complicated BECCS processes show great promise for use in energy generation and carbon intensive industrial processes like cement and steel production. In all cases the energy produced is carbon neutral, coming from and returning to the atmosphere, and carbon stored or incorporated in long-lived products is removed from the atmosphere, i.e., carbon negative.

Utilization

Of course, the value of carbon capture in reducing global warming depends on what is done with the carbon after it is captured. One option is to use it in place of carbon from fossil sources. Some uses keep the carbon out of the atmosphere for only a short time, such as synthetic jet fuel (**FIGURE 4**). Still. If the carbon was extracted from the atmosphere to begin with, using renewable energy, then the net carbon contribution when it is burned can still be zero. It becomes a way to make green energy available to industries like aviation where finding a substitute for hydrocarbon fuels is challenging.

Other uses, such as building materials, retain carbon for a very long time. Synthetic aggregates made by mineral carbonation can provide essentially permanent storage of CO_2. The carbon dioxide may come from flue gas or other pure streams. Synthetic aggregates can also be used to dispose of industrial wastes such as fly ash, steel slag, and cement kiln dust. Concrete is 60% to 80% mineral aggregate. In traditional concrete, the aggregates are mixed with water and cement. The water converts the cement to interlocking crystals that bind the concrete together. CO_2-cured concrete uses non-traditional cements that cure when mixed with CO_2. This is a mature technology that can be cost competitive with traditional concrete, though it is currently used primarily for precast concrete blocks.

Storage

Storage in geological formations, such as oil and gas reservoirs, unmineable coal seams, and deep saline reservoirs, is generally considered to be safe and secure when the storage facility is properly sited and operated. In nature, CO_2 remains trapped in these formations for millions of years. We have extensive experience storing CO_2 underground, with industrial scale projects dating back to the 1970's. Oil and gas producers use CO_2 injection in advanced recovery techniques to force the last bit of oil from a deposit. This should not be taken as license to continue emitting GHG with the thought that it can simply be put back into the ground where it came from. Though known potential geological storage capacity is large, it is not unlimited. The ultimate solution remains the reduction of emissions.

Semiconductor industry

What role does CCUS play in semiconductor manufacturing? Though we are not major direct emitters (scope 1 emissions), we do consume a lot of electrical power (scope 2 emissions). Our scope 1 emissions are roughly half our scope 2 emissions and consist mostly of non-CO_2 greenhouse gases such as PFCs, HFC, NF_3, SF_6, and N_2O. Short-term, we will realize the greatest benefits, both environmental and financial, by reducing power consumption through improvements in energy efficiency. Longer term we must continue to advocate for conversion to renewable energy sources and the decarbonization of the power grid. We must also remain vigilant in our efforts to capture and abate harmful GHG in the process exhaust stream, especially fluorinated gases, which are widely used for etching and chamber cleaning. Though the amounts may be small relative to global CO_2 emissions, F-gases are thousands of times more potent in their global warming potential, and often far longer-lived. SD

Measuring the Critical W-recess in 3D NAND

NICK KELLER, **ZHUO CHEN**, **PETER WANG**, **ROSTISLAV GRYNKO**, **TROY RIBAUDO, G. ANDREW ANTONELLI** and **YOUCHENG WANG**, Onto Innovation, Inc.

A new technique that uses a specialized via to periphery target structure overcomes limitations of existing approaches.

LAST SUMMER, THE SEMICONDUCTOR industry reached a significant milestone: one of the world's top-tier fabs had begun production of the first 3D NAND chip with more than 200 layers. The announcement was significant but not a shock. Several other fabs had been progressing toward breaking the 200-layer barrier, so reaching the milestone was not a matter of if but when.

As significant as this advance is, the high-volume manufacturing challenges of producing high-aspect ratio (HAR) 3D NAND chips are considerable. One challenge is the ability to measure the tungsten (W) recess to the bottom of a 3D NAND device following the replacement gate process. Presently, there is no in-line process control solution that can accomplish this. The reason for this is known: beyond just a few layers in the stack, the W recess becomes opaque in the ultraviolet/visible/near-infrared region, the realm of many OCD systems, after just a few layers in the HAR stack. Additionally, increased wordline slit pitch scaling further reduces the already minimal optical signal from the top of the 3D NAND structure to the bottom.

Both of these matters present a challenge for manufacturers because the ability to measure the W recess is critical. Under-etching the W gates in the recess will cause wordlines to short, while over-etching the W gates could damage cells adjacent to the wordline slit or cause a short from the wordline to the source line.

Fortunately, a new technique can overcome these W recess challenges in HAR 3D NAND devices.

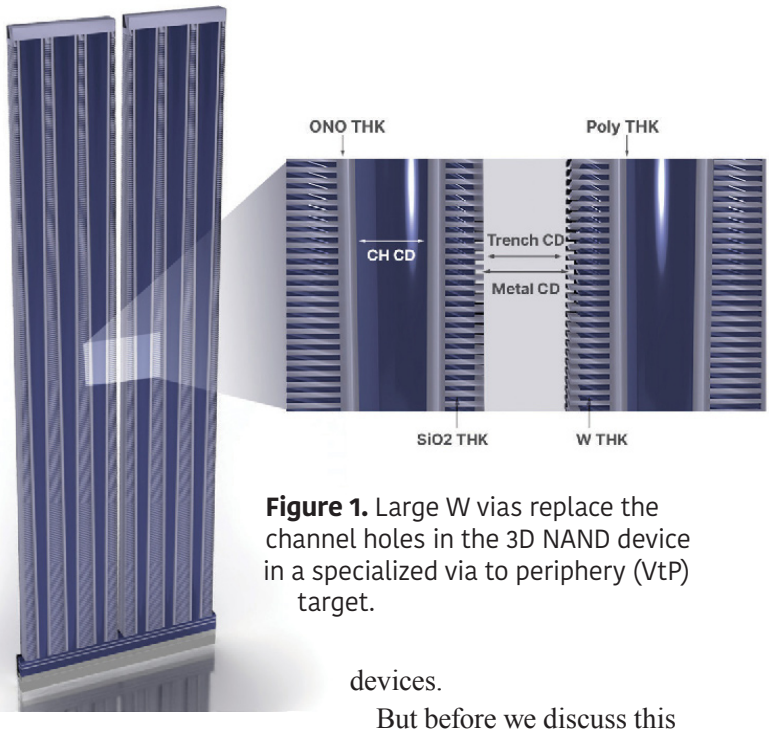


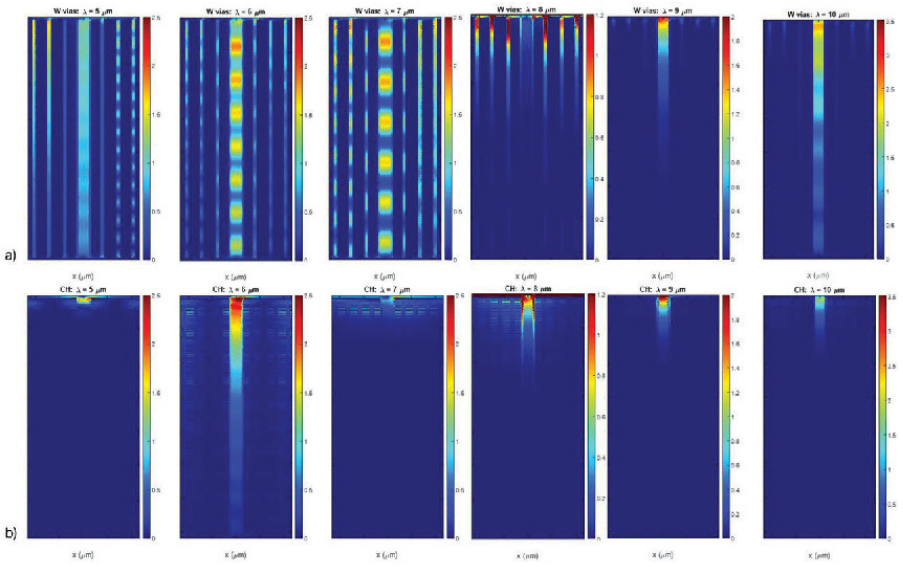
Figure 1. Large W vias replace the channel holes in the 3D NAND device in a specialized via to periphery (VtP) target.

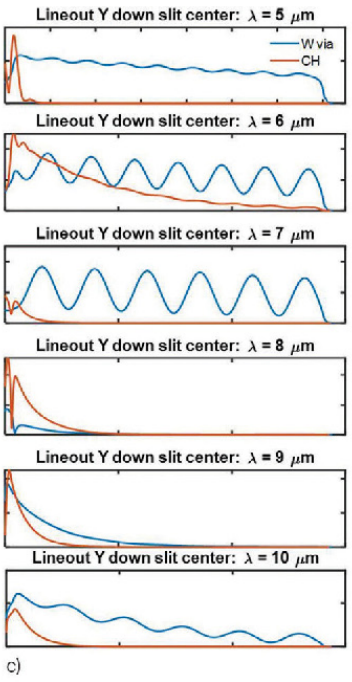
By using finite-difference time-domain (FDTD) and optical critical dimension (OCD) simulations, we were able to determine that a specially designed, design rule-compliant ellipsometry target enables mid-infrared (IR) light to penetrate through metal oxide pairs to reach the bottom of the W recess, thus enabling the measurement of the Z-profile of the W recess. Equally as important for the future of HAR devices, this data demonstrates that mid-IR light can overcome the challenge of reaching the bottom of a W recess in a 3D NAND chip with more than 200 layers, a step that might be significant in the proliferation of these devices.

But before we discuss this new technique, we need to better understand plasmonics.

Understanding plasmonics

Plasmons are collective longitudinal oscillations of electrons in a plasma. Surface plasmon polaritons exist at the interface of a metal and dielectric and can be excited by photons if they are polarized in the plane of incidence and meet two other criteria. The first criterium that must be met is that incident photons and surface plasmon polaritons need to share the same frequency and momentum [3]. To match the frequency and momentum of the photon and surface plasmon polaritons, a coupling medium, like a prism or grating, must be used because a photon





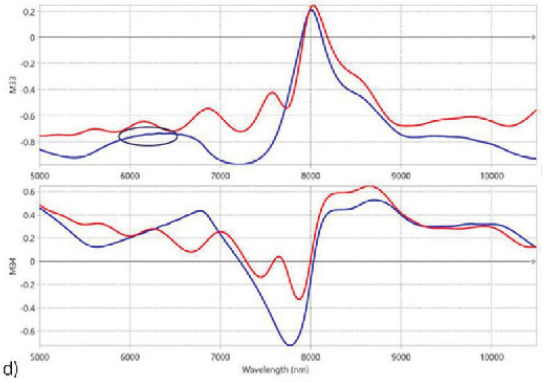


Figure 2. (a) FDTD simulations of the VtP structure from wavelengths of 5 to 10μm at wavelength increments of 1μm and (b) FDTD simulations of 3D NAND device structure. (c) Electric field amplitude in the center of the wordline slit as a function of Z for VtP structure (blue) and 3D NAND structure (orange). (d) RWCA-based spectral simulations of VtP (red) and 3D NAND device (blue) structures with circled area corresponding to highly attenuated oscillations indicating the electric field penetrates to the bottom of the slit.

in free space has less momentum than a surface plasmon polariton. The second criterium is that the two materials must have opposite signs for the real part of the dielectric permittivity. Negative dielectric permittivity can be observed in metals and doped materials when the frequency is less than the plasma frequency. The dispersion relation relates the frequency of light to the spatial wavenumber of the surface plasmon polaritons and the equation for a single metal/dielectric interface system.

Simulations and motivation for target design

The plasma frequency is defined as when the real part of the dielectric function equals zero and becomes negative for lower frequencies [5]. For W, that wavelength is 942nm. IR light lower than this frequency can excite diffraction-assisted volume plasmonic resonance, following the metal gate replacement process step, in 3D NAND structures with wordline slit pitches less than one micron, allowing mid-IR light to reach the bottom of the wordline slit in the Z direction [6].

However, as 3D NAND scales, both wordline pitch and the number of tier stack pairs increase. This causes an exponential loss of sensitivity in the W recess despite the material's ability to support plasmonic resonance. Rigorous coupled wave analysis (RCWA) simulations run in Onto Innovation's proprietary Ai Diffract™ software in mid-IR demonstrate a decrease in sensitivity to the bottom W recess versus the wordline slit pitch by showing an increase in intrinsic uncertainty [7]. Further RCWA simulations were run in the software in the mid-IR region where Onto's Aspect® system operates (5 to 10.5μm) to validate the observation that an increase in wordline slit pitch reduces sensitivity to the bottom W recess.

Through-cell contacts were introduced to connect the lower metal layers below the array to the upper metal layers which are processed after the 3D NAND array is completed. This was done in conjunction with putting the peripheral circuitry under the NAND array. These through-cell contacts are large W vias (100s of nm in diameter), while the W vias are taller than the 3D NAND stack in the vertical direction and placed in the staircase region of the die. If these large W vias replace the channel holes in the 3D NAND device in a specialized via to periphery (VtP) target (**FIGURE 1**), they can act as a plasmonic waveguide and couple IR light, or more specifically wavelengths longer than the plasma frequency, to surface plasmon polaritons that can reach the bottom of the wordline slit.

Comparison of 3D NAND array structure to VtP target

To assess the feasibility of using a VtP structure as a plasmonic waveguide, FDTD simulations were

performed using the Lumerical™ FDTD software package on a 3D NAND device structure as well as a VtP structure with the same wordline slit pitch at the W recess step. RCWA-based spectral simulations were also run from wavelengths of 5 to 10.5µm. Both structures have 2XX layer pairs, a wordline slit pitch greater than 2µm, a slit width of roughly 250nm and a W recess of roughly 50nm. While the 3D NAND device structure has channel holes of roughly 100nm diameter between the slits, the VtP structure includes three W vias lined with SiO_2 and which have diameters greater than 250nm. Of these three vias, the central via is located at the half pitch, and the outer vias are located at roughly the quarter pitch of the structure.

FIGURE 2 (a and b) shows the results of the FDTD simulations for the 3D NAND device and the VtP structure from wavelengths of 5 to10µm in increments of 1µm. Fig. 2c shows a comparison of the electric field amplitude in the center of the wordline slit as a function of Z at the same wavelengths as Figure 2a. Figure 2d shows a comparison of RCWA-based spectral simulation of M33 and M34 components of the Mueller matrix for both structures from wavelengths of 5 to 10.5µm. The FDTD simulations show the amplitude of the electric field as it interacts with the structure.

The visible asymmetry in the electric field of the VtP structure is due to the oblique incidence angle of illumination and the phase difference induced by the distance between the source and the various apertures, which are the tops of the vias and wordline slit opening.

For the VtP structure, light can couple to SPPs and propagate through the wordline slit and SiO_2 liners to the bottom poly-Si layer at wavelengths of 5, 6, 7 and 10µm. At wavelengths of 8 and 9µm, the absorption of SiO_2 is largely due to the strong Si-O bond, causing the real part of the dielectric function to become negative, so that SPPs cannot be excited. SPP oscillations in the wordline slit are visible at wavelengths of 6, 7 and 10µm, and SPP oscillations are also apparent in the SiO_2 liner surrounding the W vias at a shorter wavelength compared to the slit region due to the smaller gap distance [3]. Fig. 2c and 2d show good agreement, indicating that FDTD simulations offer an accurate physical model of how SPPs propagate in the structures. Note that the electric field propagates to the bottom of the device structure at a wavelength of 6µm (shown in Fig. 2b and 2c), and it corresponds to highly attenuated oscillations in the simulated spectra from wavelengths of 6 to 6.5µm (circled in Fig. 2d).

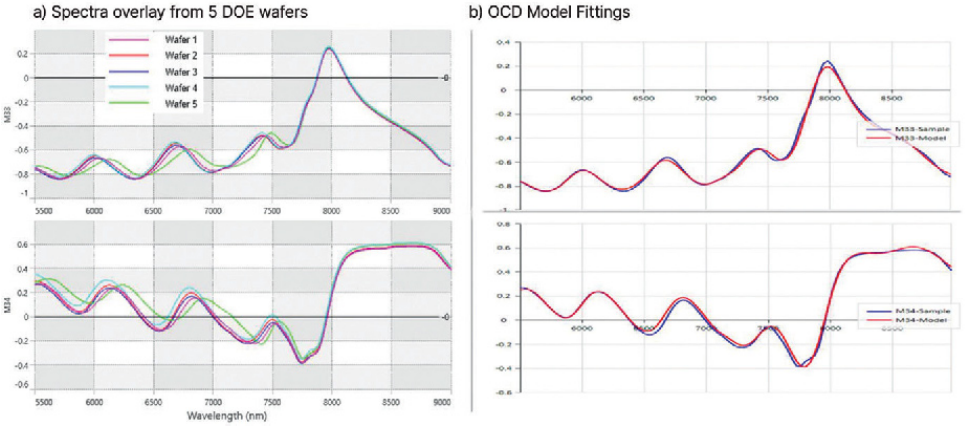


Figure 3. a) Mid-IR spectra overlay from five DOE wafers (one center site selected from each wafer) b) OCD model typical fitting.

Experimental results and analysis

The proposed VtP structure was printed and processed as a scribe line target with a 2XX tier 3D NAND device process. Five wafers were prepared with a design of experiment (DOE) at three steps: – photolithography, dry etch and wet etch recess – to create skew on the slit CD, slit etch profile and W recess depth, respectively (Table 1). The wafers were measured on a mid-IR ellipsometer to get mid-IR spectra for the development of the optical critical dimension (OCD) model. Then we measured three sites on each wafer using a destructive reference tool to determine oxide CD and metal CD profiles. For each site, we measured over 130 positions from top to bottom. The average of each 10% height zone was used to optimize the OCD model.

The spectra from the metrology system were then fit in the software. We used software embedded with machine learning to optimize the model and then used a model-guided machine learning (MGML) recipe to determine the results of the oxide CD and metal CD in the Z-profile. These results were then used to deduce the Z profile of the W recess.

Fourteen locations were measured with a destructive reference tool, and nine reference measurements were used for MGML training. The other five reference measurements were left to be blind tested to validate the robustness

Table 1. DOE on five wafers with skews from three process steps: photolithography, dry etch and wet etch recess.

WAFER	SLIT CD	SLIT ETCH	W RECESS
1	POR	POR	POR
2	POR	POR	Recess+
3	POR	POR	Recess++
4	CD+	ALT	POR
5	CD-	POR	POR

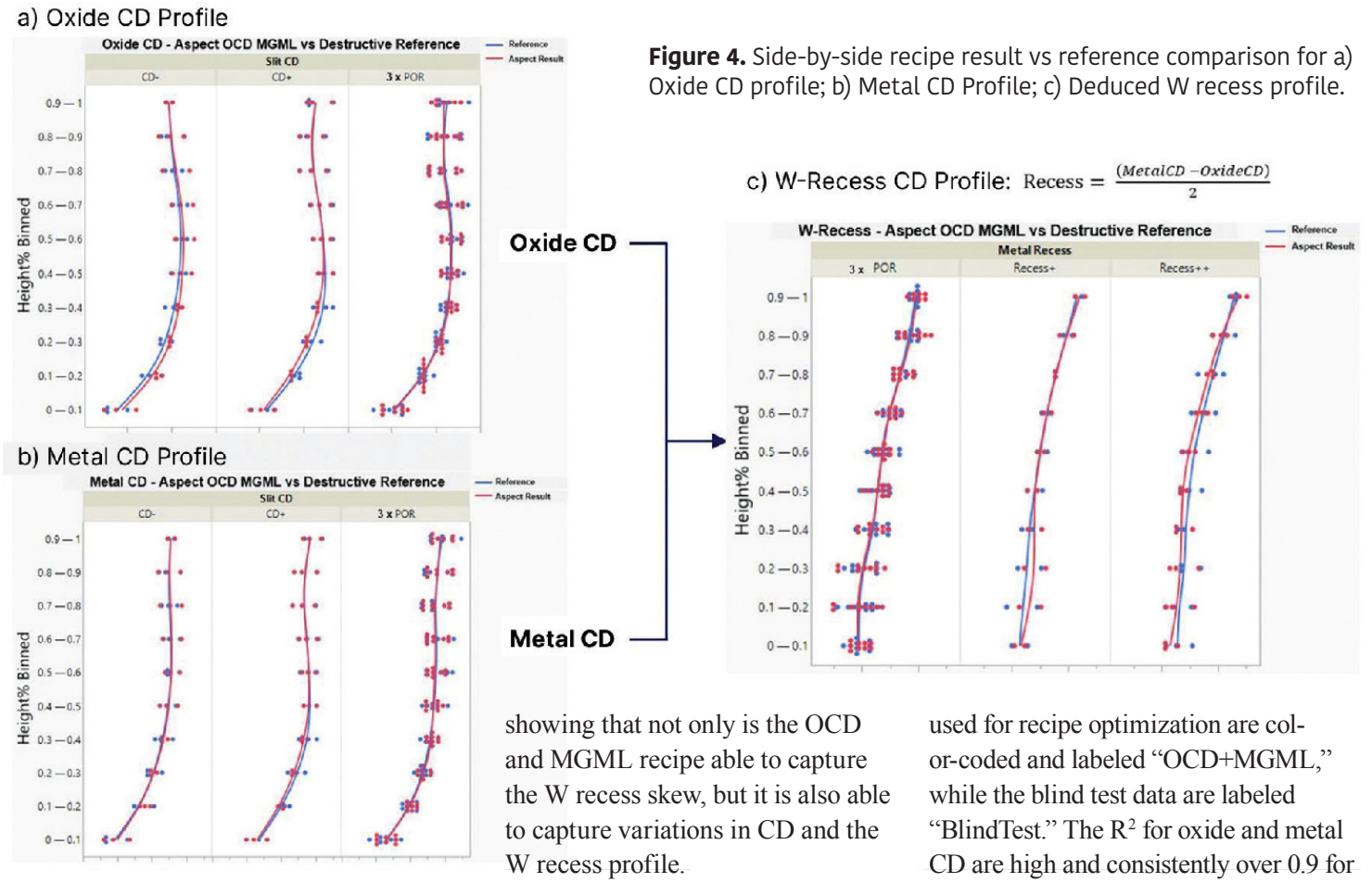


Figure 4. Side-by-side recipe result vs reference comparison for a) Oxide CD profile; b) Metal CD Profile; c) Deduced W recess profile.

of the recipe. The result of the OCD and MGML recipe is highly consistent with the references from the destructive tool, showing that not only is the OCD and MGML recipe able to capture the W recess skew, but it is also able to capture variations in CD and the W recess profile.

In **FIGURE 5**, each height zone from each site that has a reference as a data point was used to plot the recipe result versus reference correlation. Reference points used for recipe optimization are color-coded and labeled "OCD+MGML," while the blind test data are labeled "BlindTest." The R^2 for oxide and metal CD are high and consistently over 0.9 for both recipe optimization data and blind test data. For the W recess, the blind test is slightly lower, 0.79. However, it should be noted that we intentionally left recess skew edge sites for blind test as validation for robustness Continued on page 35

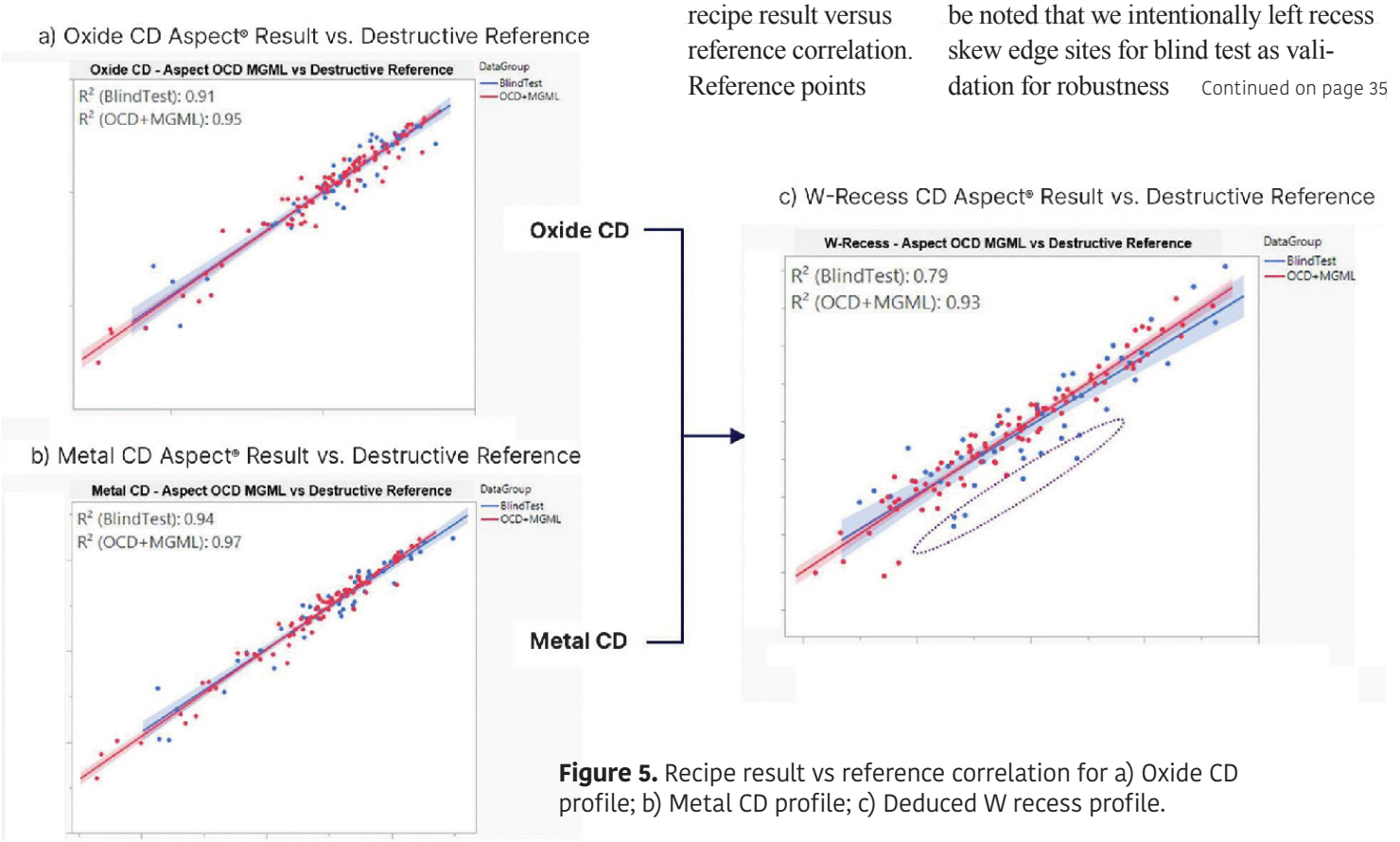


Figure 5. Recipe result vs reference correlation for a) Oxide CD profile; b) Metal CD profile; c) Deduced W recess profile.

Continued from page 27

AMD's Papermaster noted that it was just a decade ago that deep neural net started to show the promise of what accelerated computing could be. "There's no question that we're still at the earliest stages," he said. "But I am extremely hopeful because I know that if we take on the right policies and the right constraints, what we're going to see in the coming years is that this transition is going to enable us as society to be incredibly more productive. We will speed innovation with the inflection of AI powered computing."

Van den hove said it will require "moonshot" type thinking. "It will require skyrocketing collaboration across the entire value chain and across the continents, bringing together the best of the best. Just like at the time of the Moon mission in the '60s, we're faced with a seemingly insurmountable challenge. Once again, we'll have to deliver as one, building up a rocket fueled with innovation. Except now we're not just aiming at the moon, we're aiming at a better planet, a better life on planet Earth."

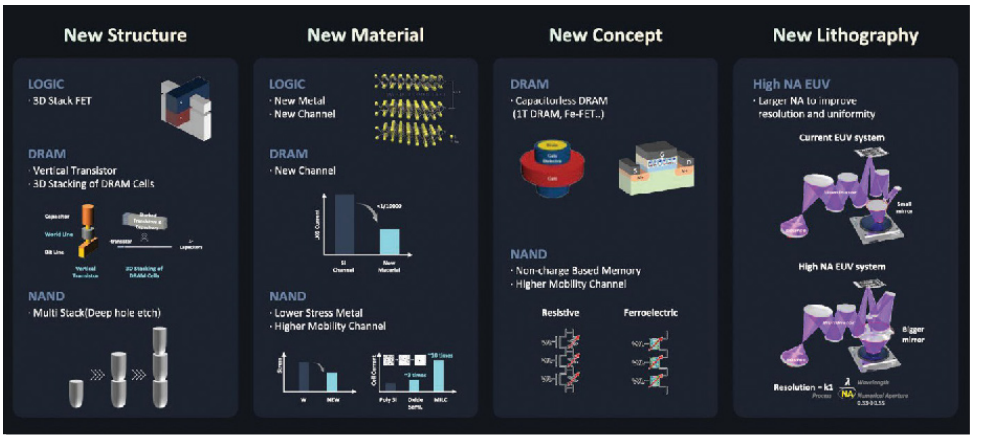


Figure 6.

Metrology Continued from page 34

to handle process variation since the edge sites have much larger process variation compared to all prior steps. The points that have larger correlation residue below the fit line were all from one site. If this edge site is excluded, the blind test R^2 will be over 0.85.

Conclusion

The tungsten recess replacement step is a key part of the 3D NAND manufacturing process. However, current metrology systems struggle to provide Z-profile measurements of this recess in HAR devices because channel holes are responsible for blocking the coupling of mid-IR light to SPPs in 3D NAND structures, which is likely caused by poly-Si channel material. The technique outlined in this article of using a VtP structure, the wordline slit pitch and model-guided machine learning provides a way this important challenge can be overcome. Equally as important, this technique demonstrates the ability to recess skew and capture CD/recess profile variation for the entire W recess in HAR devices of 200 or more layers. This ability will be required for the high-volume manufacturing of bleeding edge 3D NAND devices.

About the author

Nick Keller has been with Onto Innovation since its founding in 2019 and prior to that Nanometrics, which merged with Rudolph Technologies to become Onto, since 2007. He has held many roles in Applications and Corporate Marketing. Keller has been instrumental in developing infrared critical dimension (IRCD) technology and providing path-finding simulations in support of other new products at pre-initiation phases. He has four issued patents and over 20 publications. He received a Bachelor of Science in Electrical Engineering from the University of California Santa Barbara in 2007.

REFERENCES

1. Antonelli, G. A. et al. Ellipsometric critical dimension metrology employing mid-infrared wavelengths for high-aspect-ratio channel hole module etch processes. in Metrology, Inspection, and Process Control for Semiconductor Manufacturing XXXV (eds. Adan, O. & Robinson, J. C.) vol. 11611 116111O (SPIE, 2021).
2. Keller, N. et al. Novel inline on-device measurement of silicon nitride lateral recess post channel hole ACI with IRCD metrology. in Metrology, Inspection, and Process Control XXXVI (eds. Robinson, J. C. & Sendelbach, M. J.) vol. 12053 120530T (SPIE, 2022).
3. Maier, S. A. Surface Plasmon Polaritons at Metal/Insulator Interfaces. in Plasmonics: Fundamentals and Applications, 21–37 (Springer US, 2007). doi:10.1007/0-387-37825-1_2.
4. O'Mullane, S., Keller, N. & Diebold, A. C. Modeling ellipsometric measurement of three-dimensional structures with rigorous coupled wave analysis and finite element method simulations. Journal of Micro/Nanolithography, MEMS, and MOEMS 15, 44003 (2016).
5. Potter, K. S. & Simmons, J. H. Chapter 2 - Optical properties of conductors. in Optical Materials (Second Edition) (eds. Potter, K. S. & Simmons, J. H.) 69–100 (Elsevier, 2021). doi:https://doi.org/10.1016/B978-0-12-818642- 8.00002-8.
6. Yoon, J. W., et al. Nanophotonic identification of defects buried in three-dimensional NAND flash memory devices. Nat Electron **1**, 60–67 (2018).
7. Vagos, P., Hu, J., Liu, Z. & Rabello, S. Uncertainty and sensitivity analysis and its applications in OCD measurements. in Metrology, Inspection, and Process Control for Microlithography XXIII (eds. Allgair, J. A. & Raymond, C. J.) vol. 7272 72721N (SPIE, 2009).
8. Neutens, P., Lagae, L., Borghs, G. & Dorpe, P. van. Plasmon filters and resonators in metal-insulator-metal waveguides. Opt. Express 20, 3408–3423 (2012).

Semi-cap Supply Chain Overshoots

JOHN WEST, Senior Director of the Semiconductor Subsystems and Test Division, Yole Group.

Suppliers take advantage of the slowdown to get back on track.

SUPPLIERS OF SUBSYSTEMS, components, modules, and sub-assemblies used in semiconductor manufacturing equipment have experienced a huge surge in demand over the past four years. According to the Semiconductor Equipment Subsystems Market Monitor from Yole Intelligence, the relentless run of 15 consecutive quarters of growth, starting in Q1 2019 and ending in Q3 2022, saw quarterly revenues triple in response to unprecedented demand from OEMs and chipmakers.

This situation created a lot of pressure for suppliers to keep adding capacity and drive-up deliveries at all costs. When the market entered a period of falling orders at the end of last year, there was a rush to keep delivering products to clear backlogs before customers could cancel orders. Now that these backlogs have cleared and new orders are falling, inventories of parts held at OEMs and their suppliers are spiking upwards. Returning to normal inventory levels will not happen for a few more quarters. While the current slowdown is concerning, suppliers are taking this pause in demand as a much-needed opportunity to fix problems and get into a position ready for the next round of growth.

The challenges experienced by supply chains due to COVID-19 and de-risking driven by geopolitical tensions are well known. What needs to be fully appreciated is the excellent job the supply chain did to keep the semiconductor industry on a growth trajectory under difficult circumstances. Over this period, shortages of raw materials, people with the right skills, and the need to increase capacity created winners and losers. Key suppliers like MKS Instruments, Advanced Energy, Comet, and VAT Valve were clear winners as they met the challenges head-on. And overall, the supply chain managed to keep up and enabled equipment suppliers to double their annual sales. Yet, the data shows that OEMs compounded the stress in the supply chain by ordering more parts than they needed.

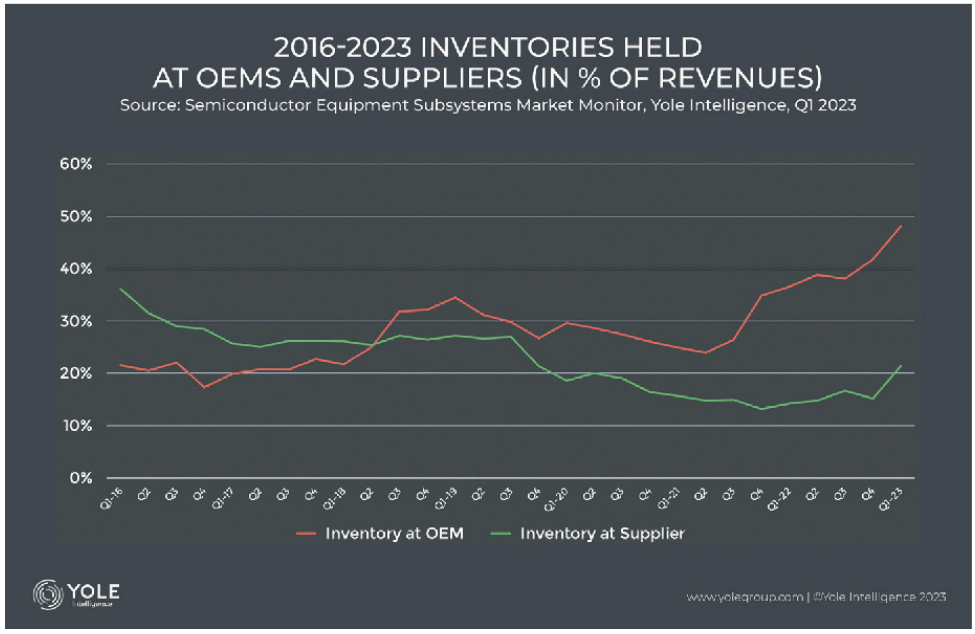


Figure 1. Inventories of finished goods held by suppliers have leapt to 20% and are expected to grow further.

Historically, OEMs maintained an inventory of parts at around 25% of current revenues, which equates to approximately three months of production. However, since the middle of 2021, it was clear that OEMs were consistently ordering more parts than required. Some of this buying behavior was understandable as the industry moved from a just-in-time to a just-in-case model, but it doesn't explain the whole story. The persistence of spot shortages for some essential parts and the unexpected and random nature of these shortages forced OEMs to keep extra inventory for additional security. Parts that proved challenging to source ranged from specialist semiconductors to raw materials such as engineering plastics and critical metals. The shift to a more cautious approach to Continued on page 59

Magnetic Resilience in Motion: Evaluating STT-MRAM Chips for Automotive Applications

SIAMAK SALIMY, Co-founder and CTO of Hprobe

A replacement memory technology -- magnetic random-access memory (MRAM) -- has been introduced and is today under adoption by car chipmakers as it answers system requirements for more advanced vehicles.

THE AUTOMOTIVE INDUSTRY IS USING memory chips in vehicles for engine control units, such as on-board instruments and sensors, cameras, advanced driver assistance system (ADAS), and mapping systems; a count of 20 to more than 100 MCUs (Micro-controller Units) can be found in all cars fabricated today from entry-level to luxury ones [1].

After about 40 years of wide scale deployment of electrical charged based memories such as embedded flash (eFlash) and Static Random Access Memory (SRAM), a revolution in embedded memory technology is happening. Indeed, MRAM, based on the electronic spin-transfer torque effect (STT), is now entering into wide scale deployment as embedded memory in MCU applications. STT-MRAM presents stellar performances by very well compromising speed, power consumption, endurance, and scalability. It has already been introduced at 22nm technological node to replace eFlash memory in CMOS and FD-SOI processes for consumer wearable and IoT products [2,3]. It has also been presented on a 16nm FinFET process by TSMC [4] and 14nm by Samsung [5]. Recently, TSMC confirmed STT-MRAM commercial deployment on 16nm FinFET in a partnership with NXP [6].

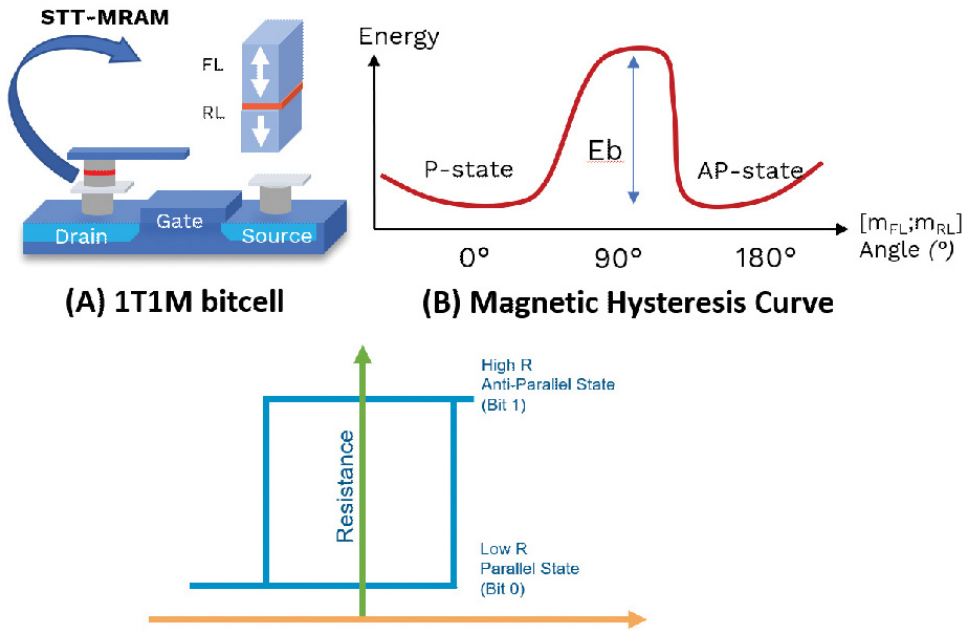


Figure 1. STT-MRAM description.

According to Objective Analysis and Coughlin Associates, in a report "*Emerging Memories Enter the Next Phase*," emerging memories have started a growth surge and should climb to become about a $44 billion market by 2032.

MRAM is by nature a magnetic memory, sensitive to certain types of magnetic fields and the adoption in automotive applications is on-going today. However, because magnetic fields can surround the STT-MRAM chip in vehicles, the chip's resilience under exposure to an external magnetic field became a critical part of the MCU test and qualification processes. Indeed, there are many different sources of magnetic fields in cars, such as magnetized wheels, alternators, electrical motors of windows lifters, and starters.

Permanent magnets are also integrated in systems for seat belt detection, door positioning, ABS and more.

The magnetic nature of STT-MRAM

STT-MRAM is a resistive memory, and the information (bit 0 or 1) is stored by the resistance value of the device. The memory bit-cell is composed of the Magnetic Tunnel Junction (MTJ) connected to a transistor or a switch selector used to route current in the MTJ for reading and writing. The MTJ is placed in the backend of line of a logic manufacturing process and positioned between two metallization layers (FIGURE 1a).

The MTJ is constructed by a stack of thin films including two magnetic ones separated by a thin enough insulator barrier to enable current tunnel effect. Each of these two magnetic layers has an intrinsic magnetization oriented in the up or down direction perpendicular to the wafer. If the two magnetic orientations are in the same direction, the MTJ resistance is in parallel state and has a low resistance value (bit '0') and if they are in opposite direction, it is in anti-parallel state and has a high resistance (bit '1'). The switching from one state to another can be triggered by temperature, current or magnetic field as long as the quantity of energy provided by the stimuli overcomes the energy barrier Eb (Fig. 1b). In the stack of thin films, a set of magnetic layers are used to give a reference orientation (called reference layer) and another layer is used to switch when writing the memory (called free layer). The free layer requires less energy to switch and can be seen as the active magnetic layer during memory operation. In the end-product operation, the writing of the MRAM by controlled switching of the free layer is made by passing an electronic current through the MTJ. At a certain threshold, the current switches the free layer by spin transfer torque.

STT-MRAM, from its magnetic nature, is sensitive to an external magnetic field. For example, to control the magnetic properties of STT-MRAM during the manufacturing process, the MRAM cells are tested under application of external magnetic field perpendicular to the MTJ surface. The magnetic hysteresis curve of the MTJ (Fig. 1c) is measured to evaluate the properties of the MTJ such as the ratio of the resistances in low and high states, the ability of changing the free layer states, the stability of pin layer…Even though in the application, STT-MRAM is operating only by use of electrical current, the level of magnetic strength to switch the free layer is normally a specified parameter to control the STT-MRAM. Depending on the stack and properties chosen for the MTJ and based on diameter in the range of 50-100nm, the free layer switching field are targeted in the range of 200 to 400mT and pin layer switching field in the range of 800mT to 1.6T, all with a direction oriented perpendicular to the MTJ surface.

Magnetic field in a vehicle

STT-MRAM deployment into automotive products requires quantifying the level of magnetic immunity of the memory as well as the level of magnetic field that are possibly present in the context of the end application.

From the literature, stray and propagating magnetic field are generally found to be lower than 5µT in the frequency range of 5 to 2kHz [7, 8]. As comparison, the Earth magnetic field is in the range of 25 - 65µT, depending on geographical position. The highest source of magnetic field in a car could originate from permanent magnets presenting a static field. Thus, we empirically evaluate the magnetic stray field over a bandwidth of DC-500Hz in different types of cars (Renault Clio, BMW i3, Peugeot 308). The magnetic field is measured at motor, alternator, doors, battery, fusebox, speakers and breakers areas and were recorded using calibrated I3C 3D Hall probe from Senis AG [9]. The levels of magnetic field measured in each considered case are illustrated in **FIGURE 2** and are shown in the range of less than 1mT at distances of more than 1 cm. The exact magnetic source position being in some cases unknown. The last right point of the graph illustrates the commonly used Neodymium permanent magnet of car phone holder. It presents here a magnetic field of 250mT at its surface and such magnets can present magnetic field in the range of amplitude for example of 100 to 500mT at the magnet surface. The field strength decreases quickly with the distance R as it is proportional to $1/R^3$[10]. As an example, for a cylindrical permanent magnet of 500mT, the magnetic field at

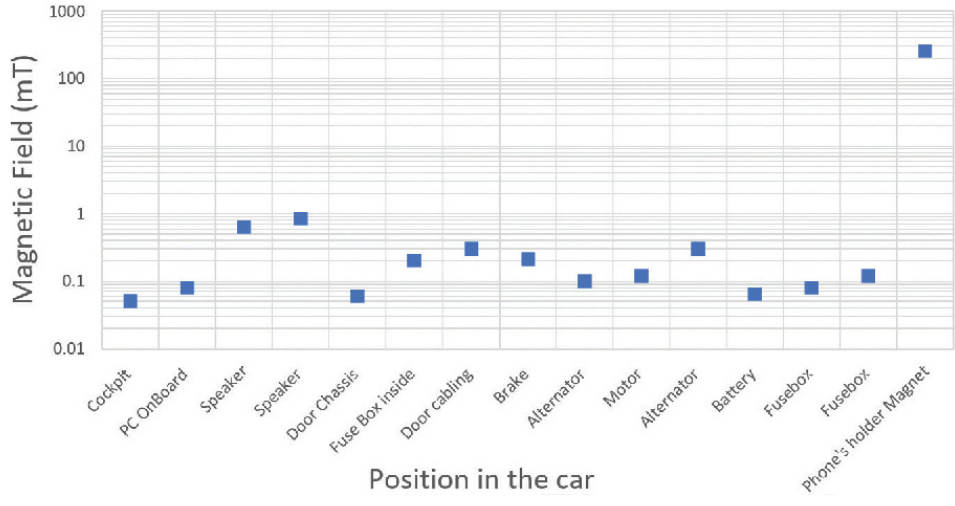


Figure 2. Stray fields in a car measured at different positions .

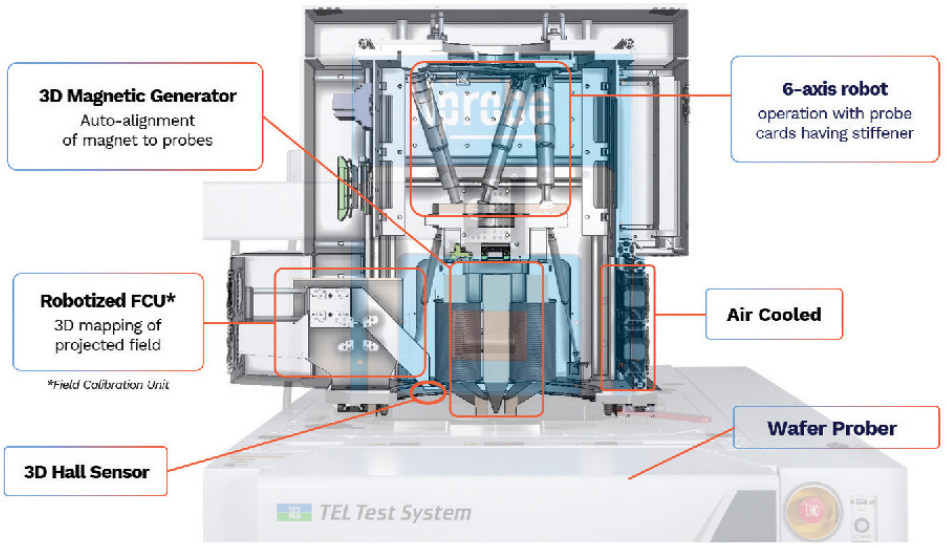


Figure 3. Magnetic test head cross for magnetic immunity tests, cross.

5mm distance is in the range of ~40mT and of 7mT at 1cm distance [11].

STT-MRAM magnetic immunity

In STT-MRAM, the probability of switching the free layer at a certain field increases with the time the MTJ is exposed. This probability is characterized by MRAM manufacturers from Bit Error Rate (BER) measurements under external field exposition. BER defines the ratio of switching error count over the total number writing/reading events requested. BER targeted levels are of less than one switching error per millions of events. The BER dependance to external stimuli like temperature and magnetic field stimuli have been studied and published by MRAM foundries (TSMC [12]], Samsung [13]], GlobalFoundries [14]). Positive temperature and external magnetic field at certain fields and angle produce an equivalent reduction of the energy barrier between the two states resulting in an increase in the BER.

The worst case being where field, and temperature are applied simultaneously. STT-MRAM manufacturers provided magnetic immunity specification of the memory as well as guidelines to specify magnetic immunity and shield the MRAM if this is required by the application [15,16]. Depending on test conditions, published data [11-15,17] report magnetic immunity levels at BER less than ≤ 1ppm with 100mT static field exposed during more than an hour at ambient temperature and in the range of 25 to 50mT at high temperature. The dependance of the BER to the field angle of the external field vector at an angle of 45° between the plane of the MRAM chip is expected to be the worst cases [12,15].

During the MRAM manufacturing process, magnetic immunity is evaluated. The wafers are heated for a specified time and exposed simultaneously to a vectorial magnetic field of controlled strength and angle. During and after this exposition, the Bit Error Rate is verified depending on the test mode: stand-by and active [14]. Stand-by immunity tests consists of applying an external field to the memory chip without executing any writing and reading operation. Active mode consists in reading or writing while applying the external field to the memory [14].

Testing STT-MRAM Magnetic Immunity

To verify magnetic immunity, Hprobe has designed a magnetic ATE using a unique magnetic generator that enables STT-MRAM magnetic immunity tests at wafer level. To cover all different possible sources of magnetic field present in the end applications, which is particularly important for Automotive chips, we designed a 3D magnetic field generator capable of applying a vectorial field with the XYZ components of direction controlled. The magnetic generator is integrated into a test head docked on the top plate of a wafer prober, as illustrated in **FIGURE 3**. The vectorial field is directly projected at wafer level to enable the evaluation of magnetic immunity before the packaging and assembly process. The test head integrates a hexapod robot used to align the generator with the electrical probe needles and the wafer under test. To calibrate and monitor the projected field on wafer, the system integrates a Field Calibration Unit (FCU) composed of a two-axis robot carrying a 3D hall sensor. The FCU can map the magnetic field in a spatial volume around the position of the test, with a grid resolution of less than 10µm. The test head is air-cooled and embeds a set of temperature sensors, controllers, and safety interlocks. The magnetic instrument is driven by an electrical rack and proprietary software interfacing with the electrical memory tester. The memory tester provides the electrical stimuli and sensing for BER tests and operates synchronously with the magnetic field projection on the wafer. The instrument is designed to reach the required level of magnetic field strength and angles to cover all types of application requirements including Automotive. The maximum out-of-plane perpendicular field is 700mT to voluntary switch the free layer, and maximum in-plane field of 200mT. The projected area on the wafer is in the range of ~3x3mm² to ~50x50mm² depending on the hardware configuration and magnetic field requirements. The projected area is sized to illuminate one or multiple STT-MRAM memory arrays depending on the magnetic immunity test case. The vectorial 3D magnetic generator is operating as a 3D magnetic arbitrary waveform generator for both static and dynamic fields projection simulating the chip exposition in the Continued on page 53

Imec Improves Memory Window of a 3D Trench Cell for Next-gen NAND Flash

MAARTEN ROSMEULEN, program director of the Storage Memory program, imec

The results mark a milestone in enabling ultrahigh-density trench-based 3D NAND Flash memories.

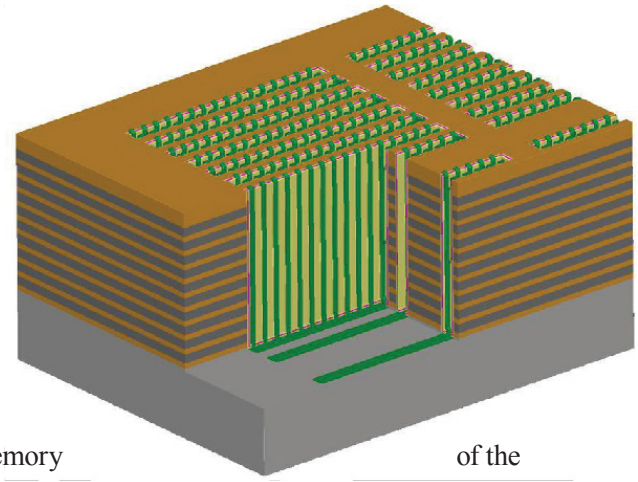
FOR SEVERAL DECADES, NAND FLASH HAS been the primary technology for low-cost and large-density data storage applications. This non-volatile memory is present in all major electronic end-use markets, such as smartphones, servers, PCs, tablets, and USB drives. In the conventional computer memory hierarchy, NAND Flash is located the furthest away from the central processing unit (CPU) and is known to be relatively inexpensive, slow, and dense compared to static random-access memory (SRAM) and dynamic RAM (DRAM).

The success of this storage technology is related to its ability to scale density and cost continuously – the main drivers for NAND Flash technology development. About every two years, the NAND Flash industry has substantially improved bit storage density, expressed in terms of increasing Gbit/mm^2.

Several technological innovations have been introduced along the road to maintain this trendline. The transition to the third dimension was arguably the most impressive innovation. In 3D NAND Flash, the memory cells are stacked to form a vertical string, and cells are addressed by horizontal word lines (**FIGURE 1**). Other notable innovations include increasing the number of bits per cell (up to four) and transitioning from the floating gate transistor to the charge trap cell for memory operation.

State-of-the-art: gate-all-around vertical channels; up to 300 word-line layers

Although not pursued by all memory makers, the charge trap cell is the base of most 3D NAND structures today. This memory cell resembles a MOSFET transistor with the addition of a small layer of silicon nitride (SiN) inserted inside the transistor's gate oxide (the oxide-nitride-oxide (ONO) stack). The SiN layer contains many charge-trapping sites that can hold an electrostatic charge. When the poly-Si gate is biased positively, electrons from the channel region tunnel through the oxide layer and get trapped in the SiN layer. This raises the threshold voltage of the transistor. The state of the cell can be measured by passing a voltage across the source/drain nodes. If current flows, the cell is in a 'no-trapped-electrons' state (corresponding to 1). Cells are in the 'trapped-electrons' (or 0) state if no current is measured.



The charge trap cell failed to be introduced in earlier 2D NAND planar configurations due to an insufficient memory window, which is measured as the difference in threshold voltage between program and erase. But in 3D NAND structures, this memory unit cell came to its full potential, thanks to a gate-all-around (GAA) vertical channel implementation method. In this GAA configuration, the gate stack completely wraps around the channel. This cylindrical geometry creates an enhanced field effect in the tunnel oxide. This leads to larger carrier injection into the trapping layers, enhancing the program/erase window.

The GAA fabrication typically starts with growing an

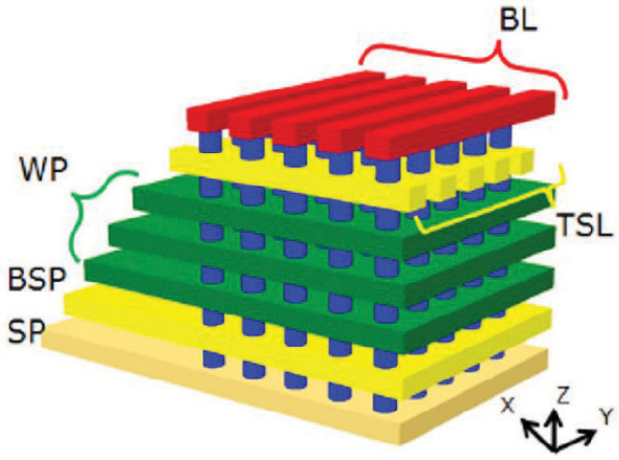


Figure 1. Representation of a typical 3D NAND Flash structure (BL=bit line; WP=word plate; BSP=bottom select plate; SP=source plate; TSL=top select line).

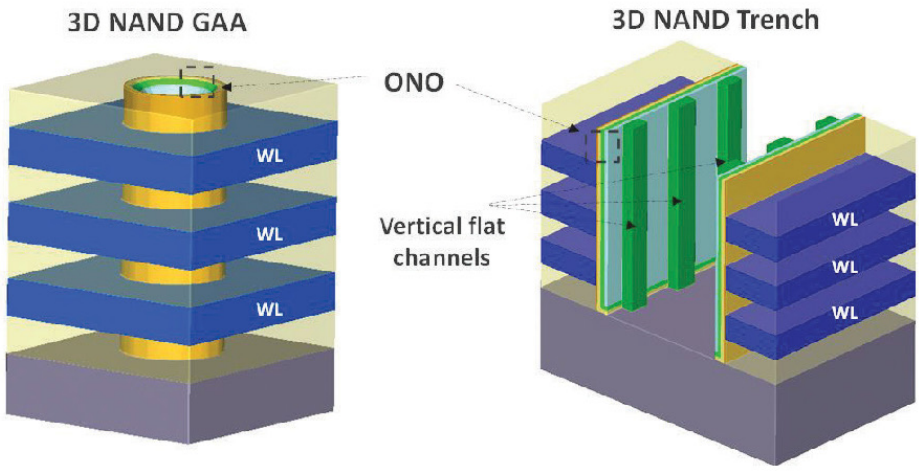


Figure 2. 3D schematics of (left) 3D NAND GAA and (right) trench devices (as presented at 2023 IMW).

oxide/word-line layer stack. Next, cylindrical holes are formed by drilling down through the stack using advanced dry etch tools. The tunneling (O) and trapping (SiN) layers and the poly-Si channel are then deposited along the sidewalls of the holes.

Recently, some major players announced the introduction of 3D-NAND-based products with up to 300 word-line layers stacked on top of each other, and this trend of increasing layers is expected to continue in the years to come. [1]

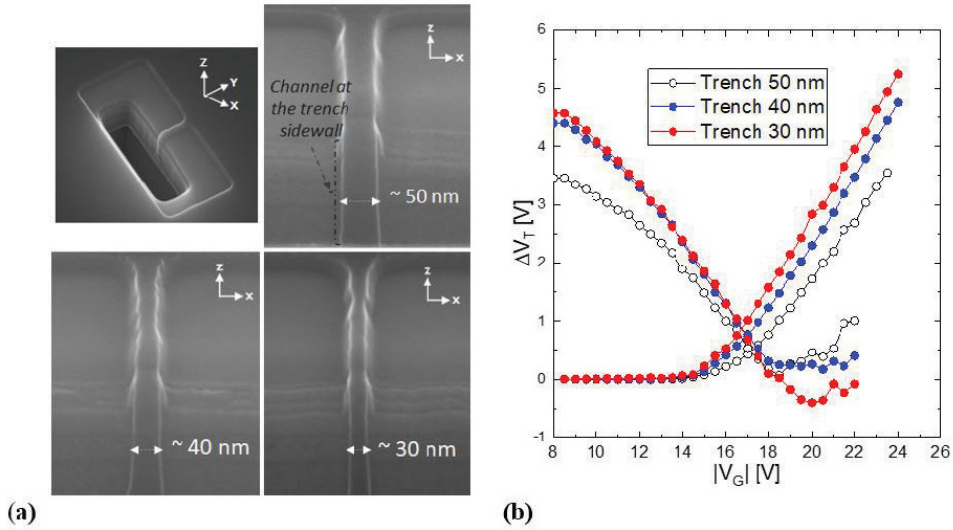


Figure 3. (a) Trench channels at different channel width; (b) program and erase characteristics, showing an improvement at smaller channel width (as presented at 2023 IMW).

Ways to further increase bit storage densities

During the present decade, memory makers will push the conventional GAA NAND roadmap to its ultimate limits. Following the most optimistic projection, the number of layers will have increased to 1,000 by the end of the decade, accounting for 100Gbit/mm^2 bit storage density [2]. This is, however, a slowdown of a few years with respect to the historical density scaling roadmap.

Increasing the number of layers introduces ever higher processing complexities and costs, challenges deposition and etch processes, and causes stress to build up inside the layers. To overcome these challenges, industry is introducing a few complementary process 'tricks' to obtain the 1,000 layers eventually. These include splitting the number of layers into two (or more) stacked tiers, further increasing the number of bits per cell, enhancing array efficiency, and reducing the GAA cell x-y pitch. There is also a trend to optimize the peripheral circuitry on a different wafer and attach it to the memory array using wafer-to-wafer bonding techniques. These innovations will, however, not be sufficient to control the growing processing costs, and therefore, an additional z-pitch scaling is pursued. Z-pitch scaling involves a reduction of the height of all materials involved in the layer stack, including word-line metals and oxides.

2030: introduction of the 3D trench cell architecture

In 2030, after GAA NAND Flash scaling has saturated, imec foresees the introduction of a new architecture to connect the charge trap cells: the trench cell architecture. With this architecture, 3D NAND moves away from the circular GAA memory cell geometry. Instead, the cells are implemented at the sidewall of a trench – resembling a planar configuration being tilted on its side – with two transistors at opposite walls of the trench (**FIGURE 2**). This next-generation NAND Flash cell architectures will not only offer the required leap in bit storage density; it is also believed to reduce costs. However, just like in the 2D planar configuration, the gate is no longer fully wrapped around the channel. Therefore, memory makers are concerned about an insufficient program/erase window.

Program and erase behavior of trench vs. GAA cell architectures

At the 2023 IEEE International Memory Workshop (2023 IMW), imec presented an experimental comparison of the memory operation of a trench cell with that of a GAA memory cell [3]. Both NAND Flash variants were processed on the same wafer, i.e., an in-house developed 3D NAND test vehicle with poly-Si gate and three word-line layers. Instead of cylindrical holes, trench features (300nm wide and 1µm long) were etched into the word-plane stack for

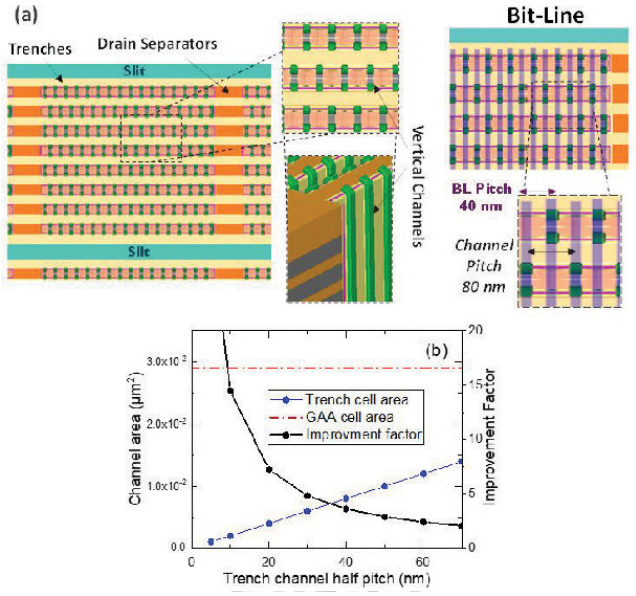


Figure 4. (a) Top view of the trench final design structure, and (b) cell density improvement factor of the trench architecture (as presented at 2023 IMW).

the trench structure. Three vertical flat poly-Si channels (with 50nm – 200nm channel width) are formed along the sidewalls of the trench, and source/drain junctions are fabricated.

Without optimization, the trench cells do not perform as well as the GAA cells. They have non-ideal program and erase efficiencies, which are reflected in the slopes and onset of the incremental step pulse programming (ISPP) and erase (ISPE) curves, respectively. This translates into a smaller program/erase window. On the erase side, the ISPE curves also show a degradation in the erase saturation level.

Toward a 5V memory window

The poor memory window can be explained by the lack of a curvature-induced field effect, which, in the GAA case, increases the carrier injection into the trapping layer. To address this shortcoming, the imec team came up with an innovative solution, i.e., reducing the channel width of the trench device. Channel width scaling is expected to enlarge the impact of the curved high-injection regions formed around the edges of the channel. In other words, at a strongly reduced channel width, the trench cell starts to resemble a GAA cell from a geometrical point of view.

On the other hand, the degradation in the erase saturation level is mainly determined by parasitic electron injection from the gate. This can be suppressed by carefully engineering the gate stack and integrating a metal gate.

Imec showed experimentally that a better memory operation could be achieved for trench devices with scaled channel width (down to 30nm), combined with an alternative high-k liner material (such as ZrO_2 or HfO_2 instead of Al_2O_3), an engineered tunneling oxide and integration of a metal gate. For most of the studied conditions, a memory window as high as 5V was demonstrated, an improvement with 2V – without impacting retention and cycling behavior (**FIGURE 3**). The team is currently working on further improving the program and erase operation.

Ultrahigh bit storage density

Having demonstrated a trench memory unit cell with good memory characteristics, the next step is to investigate potential industry-relevant integration schemes to stack a larger number of layers. Such a process flow is expected to resemble a GAA process flow, with the addition of an extra module: etching the vertical flat channel stripes at the side of the trench. Provided that a process solution can be found for this challenging etching step, imec proposed an emulation of a 3D trench process flow, with 220nm pitch trenches, each trench being 100nm wide and about 1µm long. To ensure high bit density, the flow is completed by etching 25nm wide channel stripes with an 80nm pitch (**FIGURE 4**).

From the final design structure, the trench architecture is estimated to have a three times higher cell density than the GAA reference. This is expected to further improve with channel pitch scaling. Based on these results, the 3D trench architecture can be considered as a potential breakthrough for future 3D NAND Flash memories, with bit storage densities far exceeding 100Gb/mm^2.

REFERENCES

1. https://spectrum.ieee.org/flash-memory
2. https://semiconductor.samsung.com/newsroom/tech-blog/expanding-storage-solutions-with-nand-flash-technology/
3. 'Enabling 3D NAND trench cells for scaled Flash memories', S. Rachidi et al., 2023 IEEE International Memory Workshop.

About the author

Maarten Rosmeulen received his M.Sc. degree in physics in 1993 and his M.Sc. degree in physics of micro-electronics and materials science in 1994, both from the KU Leuven, Belgium. In 2005, he received his Ph.D. in electrical engineering from the KU Leuven. Since then, he has been with imec, in Leuven, Belgium, where he has been active as an R&D engineer in process integration, semiconductor device design, and electrical device characterization for multiple internal and external projects. In 2009 he became a project leader in developing GaN-on-Silicon Light Emitting Diodes (LEDs). In 2014 he became the team leader of the Pixel Design and Testing team and has been responsible for the development of CMOS Image Sensor (CIS) technologies. In 2019 he became the program director of the Storage Memory program, the position he holds today.

Semiconductors for Automobiles: New Challenges Ahead

PETE SINGER, Editor-in-Chief

Distributed architectures, AI, wideband gap materials, chiplets and higher frequency devices are just some of the new technologies driving a new era of innovation and collaboration between car makers and chip makers.

"THE MOST POWERFUL SEMICONDUCTOR you own will not be in your gaming hardware, it will be in your car," said Stefan Hartung, Chairman of the Board of Management, Bosch, speaking at the recent imec-produced ITF World conference in Antwerp, Belgium.

Three major trends are impacting how semiconductors and other electronics in automobiles are designed and built: 1) a move from a hardware-focused approach to a software-focused approach, 2) electrification (for electric vehicles), and 3) assisted autonomous driving and control.

"Lots of new technology is starting to come in and one thing is always in common for all these three: the software content and the R&D content is exploding because that's a lot of new things to be developed," Hartung said. "At a similar pace there is a super high demand coming for reliable electronics, especially for semiconductors.

"While mobility industries once dominated as a mechanical industry, the innovation focus is absolutely shifting towards a software focus on electronics," he added. "Product differentiation is more than ever based on semiconductor functionality."

Expected changes in automotive electronics include increased use of chiplets for optimal flexibility and performance and, for electrification, increased use of wide bandgap materials, such as SiC and GaN. Hartung said there will be a "huge wave" of wideband gap devices coming. "With higher frequency and higher voltages, you can get better efficiency, longer range and less battery consumption," he said. "Semiconductors will actually unlock electrification."

Autonomous driving — or assisted driving for now — requires all kinds of sensors, high resolution, including ultra-long range sensors, radars and LiDAR, plus flicker-free video cameras, small sensors, mems, gyroscopes and accelerometers. Steering and braking will no longer be done by mechanical connections, but "by wire." 6G and over-the-air software updates are also coming.

Luc Van den hove, president and CEO of imec, also spoke at ITF World on the changing nature of automotive electronics. "Car manufacturers are facing what is probably the biggest transformation ever," he said. "Their focus has traditionally been on mechanical optimization of the car, but a much more integrated way of working will be required with hardware and software to be co-developed and co-optimized in an agile way. Much tighter collaboration between car manufacturers and chip

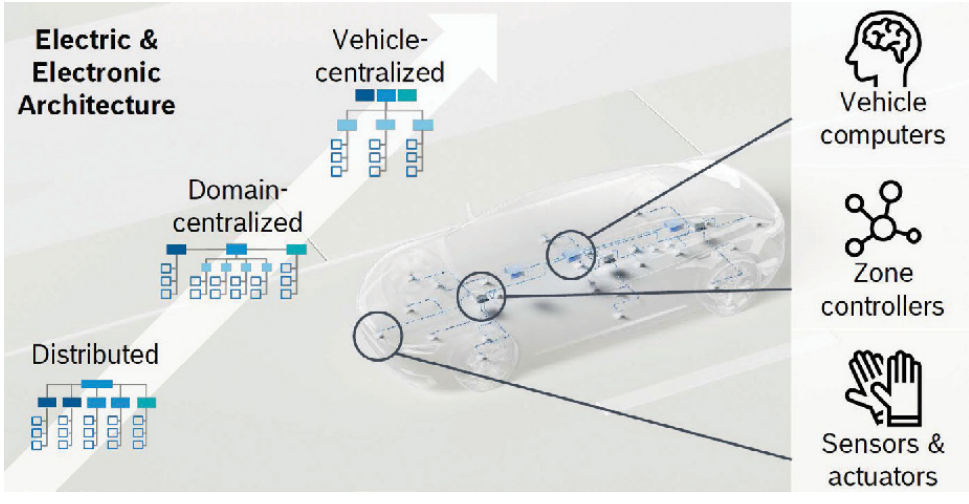


Figure 1. Vehicles of the future will have two or three vehicle computers (vehicle-centralized architecture), three or four zone controllers (domain-centralized architecture) and many edge computational devices with sensors (distributed architecture).

makers is going to be required." This collaboration should extend to the point where car manufacturers actively participate in defining the specifications and characteristics of the next generations of microchips to be used in automotive [1].

"The automotive value chain is very complicated. With innovations kind of trickling down and up along these complex value chains, which leads to an innovation process which is not compatible with the current needs," Van den hove said. "We need to bring together the ecosystem much more effectively in a network like model with the goal to accelerate the innovation across the value chain."

Rick Wallace, president and CEO of KLA Corp., also talked about automotive electronics at ITF World. "It's the fastest growing market in the semiconductor industry right now -- it's about 12.5% growth, because of the electrification and the digitization of cars," he said. "Right now, the semiconductor content of a car and the cost of that exceeds the cost of the steel and that trend is only increasing."

The compute power inside a car being delivered today is what a supercomputer was just 15 years ago, "but it has to work with heat, cold and water and ice and with sand," Wallace said. "What works in the data center doesn't work in a car and it requires us to adapt and innovate around the new requirements around this particular application."

STAR initiative

Wallace added that, on a system level, this innovation needs to evolve in way that is different than what happens in other markets because of these unique requirements of the automotive industry. "It requires collaboration on a broad scale — innovation and then execution — and it requires more players," he said.

To this end, imec and KLA Corp. recently announced their intention to establish the Semiconductor Talent and Automotive Research (STAR) initiative, focusing on developing the talent base and infrastructure necessary to accelerate advanced semiconductor applications for electrification and autonomous mobility and move the automotive industry forward.

This initiative is designed to connect automotive, semiconductor and innovation research initiatives in Europe (Belgium), United States (Michigan) and Asia (Japan). Each partner will bring relevant expertise to identify and manage programs aligned to the automotive industry, as well as talent development and recruitment.

The plan establishes a center of excellence (COE) in Michigan to formalize support for the development of the semiconductor industry workforce as well as the advancement of autonomous automotive solutions and electric vehicle research. Founding partners include General Motors, the University of Michigan, Washtenaw Community College and the Michigan Economic Development Corporation (MEDC). Specifically, the Michigan center intends to focus on:

Accelerating fundamental and advanced research for vehicle electrification and autonomous automotive solutions;

- Developing and translating of innovations in vehicle electrification and autonomous automotive technologies;
- Collaborating with learning institutions, including K-12 and vocational schools, community colleges, four-year colleges and research universities, to enable a Midwest-based skilled talent pipeline;
- Training and re-training programs to prepare the current workforce for modern chip manufacturing and assembly roles; and
- Creating a physical collaboration space, laboratory and training spaces, and incubator funding for related startups.

Distributed architectures

Hartung said the automotive industry is moving away from design approaches of the past, where electronics were in "a lot of small boxes," with each box having one functionality with one controller for such things as braking control or window control. Today, the focus is on integrating various layers to

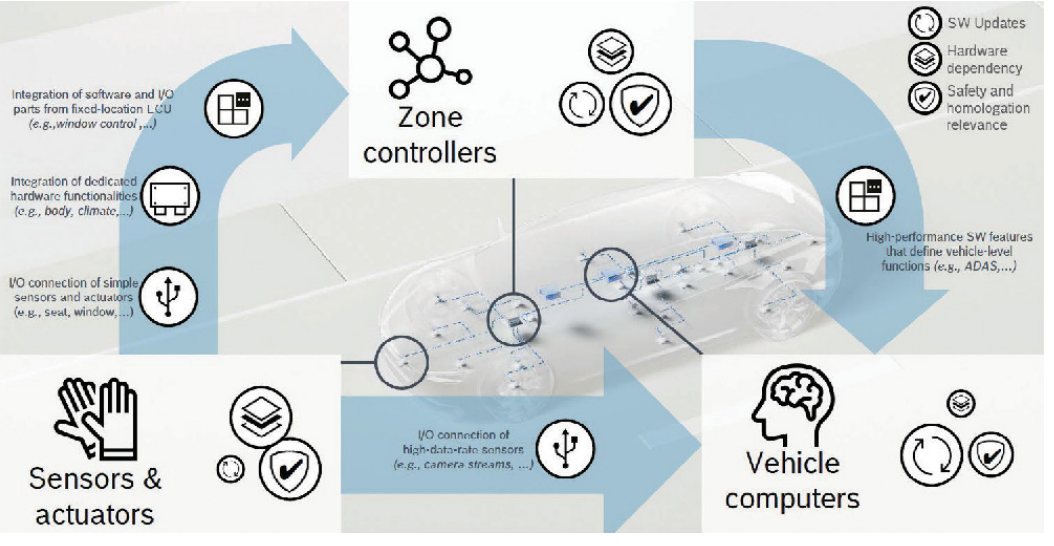


Figure 2. Chip designers would like to push as much of the edge computing as possible into the zone controller space, but face limitations related to safety regulations. Similarly, there's a desire to move functionality from the zone controller space to the vehicle computers, where it can be more easily updated over time.

implement vehicle centralized functions or zone centralized functions. "We try to take functionality from these individual boxes into zone controllers or vehicle controllers. We want to have synergies, better updatability and more functionality."

FIGURE 1 shows the three types of device classes on next generation distributed architectures: vehicle computers, zone controllers and sensors & actuators. Hartung said there might be two or three vehicle computers, three or four zone controllers and then many edge computational devices with sensors "and all kinds of high speed platforms," according to Hartung

Herbert Diess, Chairman of the Supervisory Board at Infineon Technologies said (at ITF World) the new architectures are "for sure a big opportunity to save energy. Distributed architectures will have much reduced power. We will have very few fuses anymore in those architectures, very few mechanical switches, and very much reduced wiring harnesses," he said. "This also allows the cars to become much more efficient."

Hartung described the edge level devices as being the most important, but said they will be different than those used in the past. "They will have very specific component-based functions, high speed functions, that are either sensing or acting in the vehicle," he said. These functions are normally highly hardware dependent in that there is a high-speed integration with something that is very near. "You cannot easily remove it from the edge because it's so fast. You also want to avoid too many changes in these areas because normally they are very relevant for safety regulation," he said. Reliability is paramount. "You can't allow in a vehicle any kind of fault on or loss of functionality in certain things," he said.

Hartung said designers, however, are trying to bring whatever they can from the edge to the zone (**FIGURE 2**). "When you can bring it to the zone, it's actually pragmatic to do that," he said. The same is true of moving between the zones and the vehicle. "You try to push anything from the zone up to the vehicle side," he said. "On the vehicle side, you want to have things that update frequently. You have various functionalities that are changed over their lifetime."

On a broader scale, the vehicle integrates thousands of functions that are heavily interdependent and understanding the overall system and the surrounding electronics is key. This creates a supply chain challenge. "Just-in-time supply chains orchestrate tens of thousands of parts, so supply chain resilience is a problem for the auto industry," Hartung said. "Second and third sources for anything are highly sought after, but at the same time you want to have super capable and sophisticated devices, where there might be no second source."

The case for chiplets and AI

Automakers have traditionally relied on printed circuit boards (PCBs) for most electronics, with some system-on-a-chip (SoC) solutions for high-end applications. Neither provide the kind of flexibility and performance that will be required in the future. The solution? Chiplets! "Chiplets are a better choice specifically for automotive because we can bring together the functionality of SOCs and specific components, integrate them in a high-speed environment on one chiplet, and that gives us capabilities which we didn't have before," Hartung said. Chiplets, however, require a new design approach and supply chain (**FIGURE 3**). "We need to first design a chiplet system which is fitting to the vehicle. Then we need to have chiplet elements which fit for that vehicle, which do certain tasks, computational tasks, safety, security tasks or other tasks that we'll have on the chiplet," Hartung said. "Obviously, we then have to do also the manufacturing part, which is a highly complex and very specific but possible. This can bring a new wave of high-performance computing combined with safety security requirements, combined with the rough conditions (extreme temperatures, humidity, vibration, etc.) that we have to face for automotive design."

Chiplets could also be designed with standard interfaces (**FIGURE 4**). "We need to bring together multiple players because we need lots of functionality on such a chiplet.. functionality which is not available from one semiconductor designer or one manufacturing player. We need modular and scalable architectures, and obviously what we want is a lot of new designs and a lot of new choices, higher flexibility, and mix-and-match."

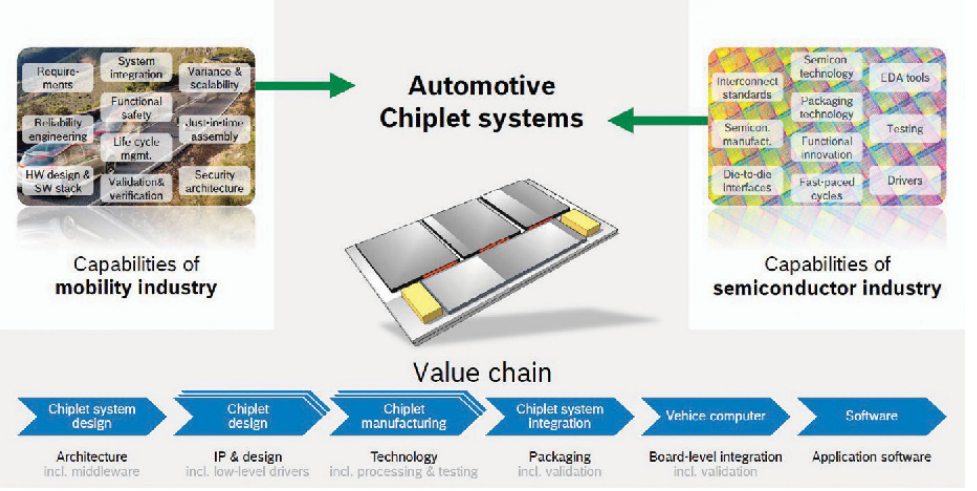


Figure 3. The capabilities of the mobility and semiconductor industries could be combined in a chiplet, boosted by a strong value chain that already exists.

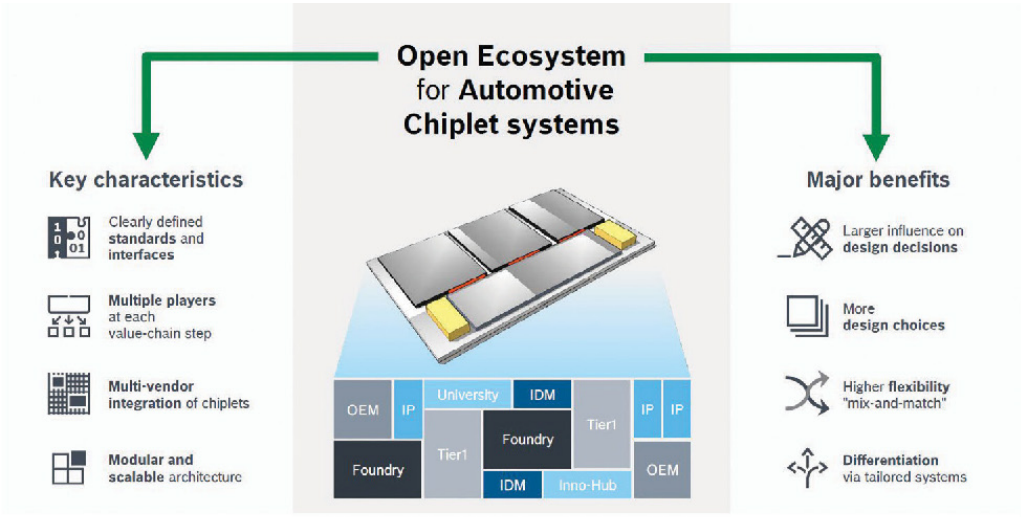


Figure 4. Chiplet design and manufacturing would greatly benefit from an open ecosystem and standardization, but this will require collaboration across a complex supply chain.

Nikolai Ardey, Excecutive Director, Volkswagen Innovation Group, also sees a need for new chip architectures and AI, and better collaboration. "We need hardware for our cars that is as flexible as a CPU and powerful as a GPU to run larger machine learning models, but they have to be as efficient as ASICs to operate within the power limitations of our vehicles," he said. "We need to collaborate with each other to examine the current trends that could lead us to this objective of flexible, powerful and efficient hardware."

A high frequency example

One recent example of the state-of-the-art in automotive components is an advanced semiconductor packaging approach from AT&S and imec that could lay the foundation for compact, cost-efficient, and high-performance 140 GHz radar & 6G mobile communications systems. At the recent International Microwave Symposium (IMS), imec and AT&S — a manufacturer of high-end printed circuit boards (PCBs) and IC substrates, presented a major step towards realizing a novel system integration approach, whereby D-Band chips and waveguides are integrated into low-cost, mass-manufacturable PCBs. This new approach paves the way for the development of compact, cost-effective, and high-performance 140 GHz (automotive) radar and 6G mobile communications systems, offering significantly lower signal loss in comparison to planar PCB lines or substrate-integrated waveguides (SIWs) in PCBs and interposers.

The need for higher bandwidths in applications such as automotive radar and upcoming 6G mobile networks is fueling the demand for higher radio frequencies. This is why the D-Band, spanning from 110 to 170 GHz, garners significant attention.

But tapping into these frequencies poses challenges, such as increased system complexity and signal attenuation. Hollow air-filled substrate-integrated waveguide (AFSIW) technology with fully metalized sidewalls offers significantly reduced signal loss compared to substrate-integrated waveguides (SIWs) that use rows of vias and the planar interconnect technology used in current communication and radar systems. Despite this intrinsic benefit, a mass-production approach for PCBs with integrated AFSIW waveguides is currently unavailable.

"To facilitate the next generation of cost-effective, high-performance radar and communication systems, millimeter-wave front-end systems that can operate at frequencies exceeding 100 GHz are necessary. However, a key concern is how these antennas and millimeter-wave ICs can be optimally connected. We believe a key part of the solution is to utilize AFSIW technology. Unfortunately, such waveguides are expensive and hard to integrate into multi-layer PCBs or interposers. Imec and AT&S have partnered to address this challenge. Together we demonstrate an integrated front-end combining imec's 140 GHz radar ICs in combination with a novel packaging/module concept developed by AT&S. The demonstration of PCB-integrated AFSIWs is an important breakthrough towards achieving this end goal," said Ilja Ocket, program manager at imec.

Conclusion

Perhaps the greatest challenge in developing new electronics for automotive applications is the timeframe. "To launch a next generation of architecture of chip design/compute platform, and set up a new industrialization process, then we are dealing with a time to market of five to seven years," said VW's Ardey. "The dilemma becomes obvious.. it's almost impossible to anticipate what we will see in about seven years at the end of this decade.

"It's clear that we are facing quite an uncertainty starting into this next generation of hardware. If car OEMs are working directly with hardware designers, that can yield several benefits: Optimized integration, customization, faster innovation, cost efficiency and enhanced safety," Ardey said.

REFERENCES

1. S. Latré and B. Placklé, "Enhanced Collaboration between Car Manufacturers and Semiconductor Industry Promises Significant Advancements in Road Safety," https://www.imec-int.com/en/articles/enhanced-collaboration-between-car-manufacturers-and-semiconductor-industry-promises.

SCREEN Brings Patterning Solutions for Advanced Lithography

HAROLD STOKES, SCREEN

In this work, we share recent activity on our DT-3000 lithography track which demonstrates the use of our patented software to improve targeted spatial locations on the wafer to provide more uniform features across the surface. We also describe an advanced integrated backside cleaning (BSC) module which allows the effective removal of an array of contaminants from the wafer backside before they are either a) transferred to subsequent processing tools such as the stepper/scanner or b) result in areas of the wafer which are out of focus on the wafer frontside.

THERE ARE MANIFOLD IMPERFECTIONS that may potentially compromise the integrity of lithographic features on today's production wafers, and as such, final yield is critically reliant on pattern uniformity. As the industry nears the introduction of high NA EUV lithography, performance metrics such as depth of focus, overlay, and CD uniformity have become increasingly stringent. CD uniformity across the entire wafer surface is dependent on meticulously controlled coating, exposure and developing processes. Part of the coating and developing process on a lithography track is the curing of the film after coating (known as PAB) and the consistent uniform application of the bake immediately following the exposure process (known as PEB). The uniformity of lithographic features across the surface of the wafer is commonly monitored after the develop process (ADI), and it is widely accepted and understood that any identified nonuniformities can potentially be improved with either lithography process tuning or corresponding etch process tuning.

SCREEN has developed proprietary software that applies a corrective model to suggest offset changes for a multi-zone plate to address nonuniformities that are typically encountered in a lithographic process. Recommended changes are then applied to the hardware to create an overall improved wafer CD uniformity. Values that are provided by the modeling software are targeted to the temperature zones of the plate to allow compensation for nonuniformities that might exist due to other process steps in the total wafer flow in the cleanroom.

Recommended offsets are dependent on several factors, and one of these is the coated material being used to form the latent image that has been exposed onto the wafer. Different photoresists have unique bake sensitivities that impact the modeling results. Quantitative information on photoresist bake sensitivity, which is required for modeling, can often be obtained directly from chemical suppliers, or it can be derived empirically from iterative exposure tests which are performed at different PEB temperatures. The specific photoresist sensitivity is added into the model so that the recommended changes will be accurate. Insufficient spatial data will result in faulty feedback from the software, therefore, properly planned measurement data is required to ensure that all

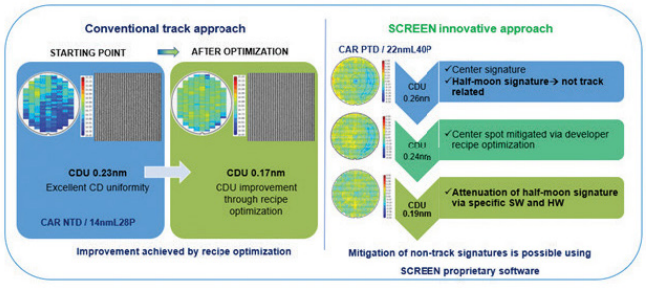


Figure 1. Conventional approach using recipe tuning alone (left), and with additional benefit of plate tuning software (right).

zones are sufficiently represented in the model.

To ensure integrity in the process it is also imperative that all process parameters are well-defined and consistent — measurement hardware and recipes; lithography processing parameters such as exposure dose and focus; film coating thickness, and baking temperatures. If any of these parameters change during the modeling exercise, then the data integrity will be compromised. Since individual hardware chambers have a unique fingerprint, such customized tuning on a plate-by-plate basis is required. Consequently, the optimal offsets for any given plate are unique to that specific chamber, requiring separate isolated tuning for each plate in the tool.

Typically, the initial modeling data input is from standard bake offsets. These standard offsets reflect those in which the thermal uniformity of the baking surface is optimized to achieve the lowest temperature range. Starting with these offsets, a customized bake recipe can be applied to the track which will allow isolation of the improvement process for the targeted material. This recipe then allows parallel improvements to be made in conjunction with the modeling software. After processing and measurement are complete the metrology data and coordinates are added into the model. The predictive modeling then supplies updated and improved offsets for each plate zone. The updated offsets are then added to the customized bake recipe for a second iteration of the test, and wafer processing is repeated with a new test wafer, measured as before, and measurement data acquired from the updated offsets is then added to the model. This iterative process results in PEB plate settings that are optimized not for the uniformity of the plate, but for the uniformity of the lithographic pattern on the wafer.

We at SCREEN have demonstrated this CD uniformity improvement capability using our DT-3000 track and an EUV exposure system to pattern 22nm lines with 40nm pitch. As shown in **FIGURE 1**, the data demonstrates a 3-sigma improvement near 20% across a 300mm wafer. After successfully applying this iterative approach, the improvement consistency can be verified with successive identical processing runs as shown in **FIGURE 2**.

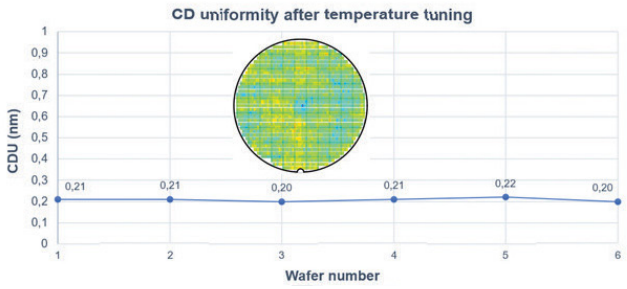


Figure 2. CD uniformity of successive wafers after optimization with modeling.

Advanced lithographic patterning depends on the ability of a given hardware set to transfer a design from photomask to a photosensitive film which, typically, has been coated onto a wafer. Countless texts and publications have described this process in detail, but as semiconductor technology continues to advance, this process is becoming increasingly challenging. Traditional obstacles such as process timing, bake uniformity, and coating uniformity have been easily routed, giving rise to issues which are currently much more challenging, such as wafer quality, stochastic defects, and backside contamination. Wafer suppliers and various process improvements have led the industry's drive to improve issues with wafer quality and warpage. Chemical suppliers and independent researchers have similarly labored to address stochastic defects. Backside contamination on wafers has been clearly identified as a hindrance to frontside imaging quality for advanced lithography, and SCREEN currently provides multiple hardware platforms and processing modules that are specifically designed to address these challenges in HVM. As designs and printed features continue to shrink, the allowable depth of focus for successful printing also continues to shrink. Consequently, previously negligible backside defects are now becoming conspicuously detrimental because the presence of defects on the back of the wafer causes corresponding front side distortions when the wafer is clamped for exposure. This distortion causes areas that are out of focus, resulting in incorrect feature sizes and pattern failures; wafer backside defects also carry the explicit risk of being transferred to the chuck of the exposure system. When defects are transferred to an exposure chuck, they commonly distort the focal plane of successive wafers in that same location when they are moved into the tool for processing. Routinely, the industry standard for addressing this problem has been targeted cleaning steps on the back of the wafer using dedicated cleanroom hardware; recently, track hardware has included cleaning modules just prior to the exposure step. However, SCREEN has leveraged cleaning expertise to further address this issue by implementing a dual-arm cleaning unit that allows customized cleaning for different contamination types, defect sizes, and defect rigidity. By demonstrating superior defect removal on the wafer back

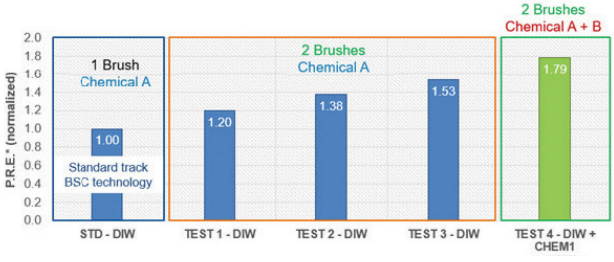


Figure 3. PRE of standard single-arm clean, dual-arm clean with only water, and dual-arm clean with optional aqueous cleaning chemistry (CHEM1).

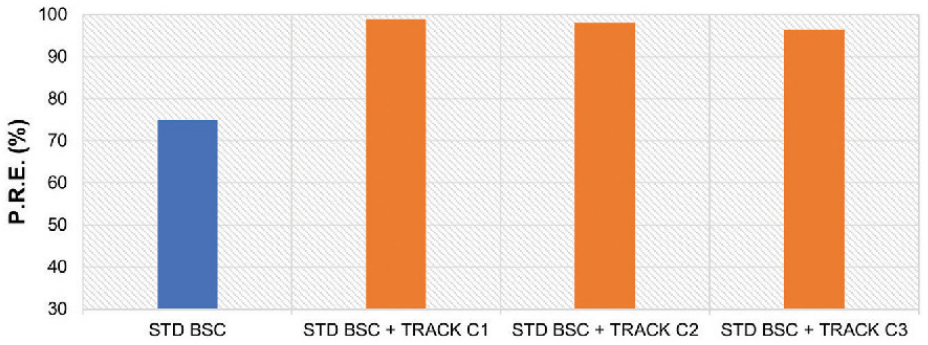


Figure 4. PRE for defects deposited with an industry standard CVD system.

side immediately before exposure, we have once again provided an industry solution that improves wafer front side patterning quality and overall CD uniformity.

Recently, we compared a standard single-arm system against our advanced dual-arm system, which can be operated with only water as a cleaning agent, or it can be combined with aqueous cleaning chemistry when required. Test results, shown in **FIGURE 3**, clearly validated our expectation that the dual-arm system is superior to the traditional single-arm system, and aqueous cleaning chemistry further improved the potential of the module. In a similar exercise, our researchers successfully demonstrated the additional benefit of using the dual-arm system to remove representative process contamination from the back of a wafer. Test wafers were cycled through an industry standard CVD system, and then patterned on an EUV exposure tool with different concentrations of aqueous cleaning chemistry used to compare against a standard pre-litho clean. **FIGURE 4** shows that the particle removal efficiency (PRE) can be improved by approximately 25% versus standard cleans. C1, C2, and C3 indicate different concentrations of aqueous cleaning chemistry.

Defect removal was verified with an electron microscope as shown in **FIGURE 5**. Below, five different sample defect positions are shown of the same wafer after standard cleaning (STD BSC) and after a subsequent dual-arm clean (STD BSC + TRACK C1). The improvement is irrefutable. As can be reasoned, such a cleaning process does not allow for the removal of pits or similar recessed areas in the substrate after the removal of partially embedded defects.

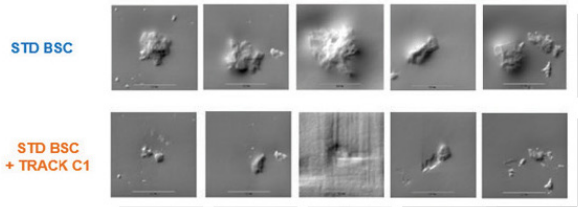


Figure 5. SEM images of corresponding defect locations after cleaning with a standard single-arm clean (top), and again after cleaning with our dual-arm system (bottom).

Phase shift response (PSR) metrology provided the most direct correlation between backside contamination and front side patterning quality. This technique allows researchers to identify locations of pattern contrast deterioration that can then be correlated to front side patten failures. When comparing backside defect locations with corresponding front side positions, the SEM images show that pattern loss and/or failure is present in these same locations. **FIGURE 6** shows the location of several backside defects and the front side wafer pattern at those same 4 locations. By clearly establishing the correlation between backside defects and front side patterning distortions, we now logically conclude that the improved performance of the dual-arm system delivers improved front side pattern quality.

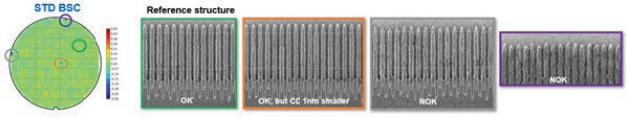


Figure 6. Relative contrast map of the wafer submitted to a standard BSC process after subtraction of intra-field signature, with the corresponding top-down SEM images collected at the positions where contrast deterioration was identified.

Developing these advanced lithography modules and options for our customers has allowed us to be more than just an equipment provider. SCREEN researchers are actively engaged in bringing solutions to market. In addition to the independent solutions discussed above, we also take a holistic approach to the development of new technology. By partnering with similarly motivated participants such as chemical suppliers we have been able to deliver solutions that could not have been implemented by a single contributor. SCREEN continues its efforts to provide world-class lithography solutions to customers and partners across the globe—and we're succeeding. Our solutions allow customers to achieve their fab targets with higher yield, higher productivity, and lower cost of ownership. SCREEN is actively engaged in multiple sustainability initiatives, and we are committed to reducing our carbon footprint to meet the requirements of modern manufacturing while providing these solutions to the industry. Production of our diverse product portfolio is supported by our new and upcoming S^3 1-4 production facilities located in Hikone, Japan. SD

Details on technical content can be directed to: imec program manager, Andreia Santos. a.santos@screen-spe.eu, lithography R&D manager, Dr. Masahiko Harumoto. m.harumoto@screen.co.jp.

Advanced Packaging: Fueling the Next Era of Semiconductor Innovation

BILAL HACHEMI, Technology & Market Analyst, Advanced Packaging at Yole Intelligence.

Advanced packaging technologies are well-positioned to address the evolving demands of high-performance computing, AI and 5G.

THE ADVANCED PACKAGING (AP) INDUSTRY is experiencing a fascinating phase of significant advancements. As Moore's Law slows down and the development of devices below 2nm nodes gains substantial R&D investments from industry giants like TSMC, Intel, and Samsung, advanced packaging has emerged as a valuable avenue to enhance product value. This approach offers advantages in both scaling and functional roadmaps, addressing the need for improved integration, the era of more-than-Moore, and the influential drivers of AI, 5G, and high-performance computing (HPC). Electronic hardware requires efficient computing power, high speed and bandwidth, low latency, low power consumption, increased functionalities, memory, system-level integration, and cost-effectiveness to support these requirements. Advanced packaging technologies are well-positioned to enable these diverse performance requirements and complex heterogeneous integration needs, making it an opportune time for businesses operating in various packaging platforms to thrive.

Despite the economic slowdown experienced by the semiconductor industry in 2022, resulting in a decline of -6%, advanced packaging exhibited resilience and achieved remarkable revenue of $44.3 billion. This demonstrates the sustained demand and importance of advanced packaging, defying prevailing macroeconomic trends.

Recognizing the rising consumer expectations for smaller, faster, affordable, and power-efficient devices, various players in the semiconductor industry, including OSATs, IDMs, and foundries, have acknowledged the significance of advanced packaging. These companies have strategically positioned themselves as key contributors to technological advancements by specializing in advanced packaging techniques. They leverage their unique business models to drive innovation and meet the market's growing demands.

A notable value transfer is from the front-end to the back-end processes. This shift is primarily driven by the increasing significance of advanced packaging, which is causing a transformation in the packaging and assembly business traditionally dominated by OSATs. Major players — such as Intel

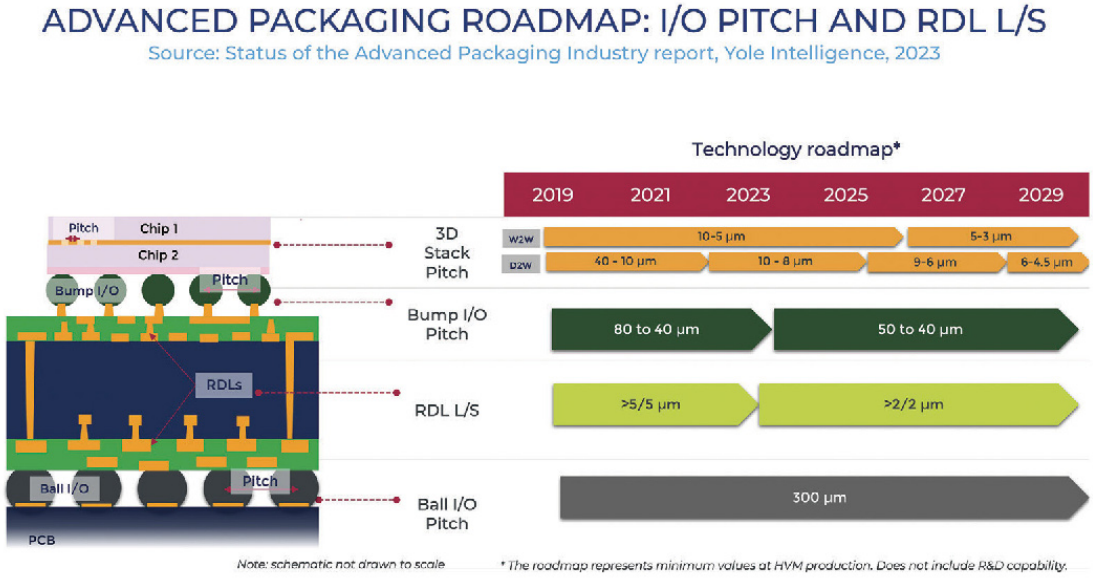


Figure 1. I/O pitch and 3D Stack Pitch roadmaps for advanced packaging.

——are strategically expanding their business models to capitalize on this market and are gradually encroaching on the market share of OSATs. These semiconductor giants actively participate in the advanced packaging segment by leveraging their strong front-end capabilities.

The AP landscape has evolved, transitioning from a package substrate platform to silicon. This shift presents substantial opportunities for industry leaders such as TSMC, Intel, and Samsung to demonstrate their prowess as innovators of new AP technologies. Among these players, TSMC has notably emerged as a frontrunner, spearheading the development of cutting-edge advanced packaging platforms ranging from fan-out (InFO) to 2.5D silicon interposer (CoWoS) and 3D System-on-Integrated Chips (SoIC).

With their capabilities, these players are driving advancements in the AP market, ushering in a new era of innovation and pushing the boundaries of what is possible in packaging technology. Their contributions are instrumental in shaping the future of advanced packaging and positioning them as key influencers in this rapidly evolving market segment.

Advanced packaging plays a role in overcoming the scaling gap between the die and PCBs. While the scaling roadmap faces challenges (**FIGURE 1**), with only three significant players remaining and a slower pace of progress, the functional roadmap, fueled by chiplets and heterogeneous integration, has gained prominence. AP technology strongly supports this functional roadmap, which adds value to semiconductor products by incorporating new functionalities and maintaining or enhancing performance while reducing costs.

To cater to consumer needs and address the demand for heterogeneous integration, multi-die packaging solutions like system-in-packages (SiPs) are being developed. These solutions offer improved functional performance and a faster time-to-market. However, this pressures AP players to innovate in processes, materials, and equipment for advanced packaging. Chiplet configurations, obtained by disaggregating or duplicating monolithic SOCs, reduce costs and time-to-market while providing flexibility in combining different FE nodes. They find increasing adoption in the server, PC, smartphone, and automotive markets. The lack of standardized die-to-die interconnect/interface solutions poses challenges, but efforts like the UCIe consortium aim to address this. Key players such as AMD, Intel, and Apple are driving chiplet solutions with investments and activities focused on the USA, China, and Europe. The industry is actively working on resolving chiplet challenges and developing open, off-the-shelf solutions through initiatives like the "Open Compute Project" and "Universal Chiplet Interconnect Express." This collaboration aims to create a universal interconnect solution while accelerating advancements in substrate manufacturing, package assembly, and test engineering.

Investments in next-generation manufacturing tools, particularly die-to-die and die-to-wafer hybrid bonding (Cu-Cu direct bonding), are essential for enabling 2.5D/3D stacking and heterogeneous integration. These advanced bonding technologies offer ultra-fine pitch interconnects, eliminating the need for underfill, UBM (Under Bump Metallurgy), and solder plating, resulting in significant form factor reduction. Key players in the industry, including TSMC with their System-on-Integrated-Chips (SoIC) platform, have made strides in developing hybrid bonding technology in-house. Others, such as SK hynix and UMC, have licensed Adeia's hybrid bonding technology. Material development is also a priority, with a focus on new dielectric materials, mold compounds, underfill materials, and solder interconnects to meet the demanding requirements of next-generation hardware. These investments and advancements in hybrid bonding and material development are pivotal for realizing 2.5D/3D stacking, heterogeneous integration, miniaturization, higher integration density, and improved signal performance. The collective efforts of industry leaders, research institutions, and material suppliers are driving exciting developments in 3D chiplet system architectures, design standards, and electrical test standards, thereby shaping the future of manufacturing highly advanced and compact electronic devices.

Moreover, achieving breakthroughs in scaling package features necessitates a sense of urgency from key suppliers within the semiconductor packaging industry. Developing advanced packaging technologies and materials will be crucial in driving the advanced packaging market's overall growth and meeting the semiconductor industry's evolving needs.

A closer look at the financials of the top 30 OSATs in 2022 reveals a notable disparity between the leading ten companies and the rest of the pack (Top 10 ranking in **FIGURE 2**). These top OSATs, backed by substantial investments, have created a significant performance gap compared to their counterparts. The companies positioned towards the end of the ranking face higher risks, particularly if they need differentiated technology or IP that can serve as an exit strategy through mergers or acquisitions.

In 2022, ASE, a renowned Mega-OSAT, emerged as the frontrunner in the rankings with an impressive revenue of **$11.9 billion** (excluding USI and including SPIL). ASE also had considerable capital expenditure, with approximately $1.7 billion invested during the year. This is nearly twice that of the next largest OSAT, Amkor, which invested **$908 million**. Bridging the gap between these players will

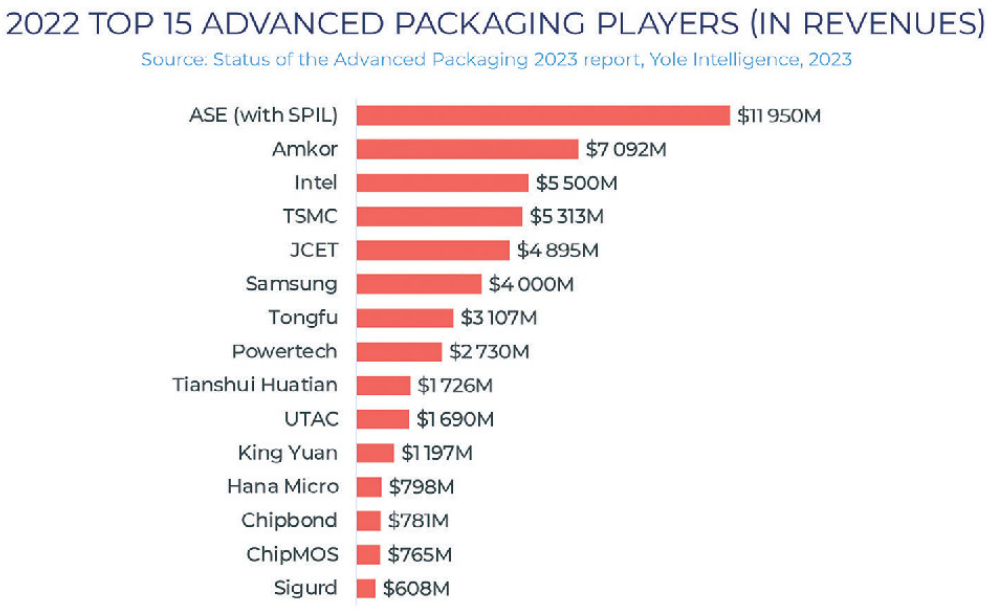


Figure 2. Advanced packaging TOP 10 Ranking in 2022 (Revenues in $M).

undoubtedly be a challenging endeavor. TSMC reclaimed the fourth position in the rankings, after Intel in the third position, reporting revenue of approximately **$5.3 billion and $5.5 billion, respectively,** for 2022. TSMC's success can be attributed to its strengthened offerings in CoWoS, InFO, and SoIC, developed for High-Performance Computing (HPC) and 5G.

Taiwanese companies continue to dominate the Advanced packaging market, commanding a significant 43% market share in 2022, surpassing other countries. ASE's exceptional performance propelled the country's position to new heights, achieving record-breaking revenue and demonstrating enhanced capabilities. The total revenue generated by AP players reached around **$60 billion** in 2022, with ASE alone contributing almost 20% of the total. Notably, the top 10 players collectively accounted for 80% of the overall AP revenue, underscoring their dominance in the industry.

In 2022, OSATs' research and development (R&D) spending decreased slightly to $1.31 billion from $1.58 billion in 2021. A significant disparity exists in R&D investment between the top six players, each exceeding $100 million, and the remaining contenders. ASE stands out as the sole player that allocated over **$450 million** to R&D initiatives. Sustaining a competitive edge in the market necessitates continuous technological innovation. In the long run, players with smaller R&D investments may need help to maintain their viability. Consequently, they are left with limited options: either increase their R&D investment or be prepared for potential acquisitions or mergers.

According to our forecasts at Yole Intelligence, the revenue from Advanced packaging is set to exceed that of traditional packaging by 2025 (**FIGURE 3**). In 2022, the total IC packaging market reached $95 billion. Within this, AP accounted for $44 billion and is projected to exhibit a CAGR of 10% from 2022 to 2028, reaching $78.6 billion by 2028. Concurrently, the traditional packaging market is expected to grow at a CAGR of 4.15% during the same period, while the overall packaging market is forecasted to achieve a CAGR of 7.1%, resulting in values of $64.7 billion and $143 billion, respectively, by 2028.

Despite the semiconductor industry experiencing a 6% year-on-year decline in 2022, attributable to factors such as economic slowdown, increasing inflation, conflicts, and inventory failures, the Advanced packaging market saw growth of 10% during the same period. The AP market is anticipated to continue growing in 2023, with an estimated year-on-year

Continued on page 55

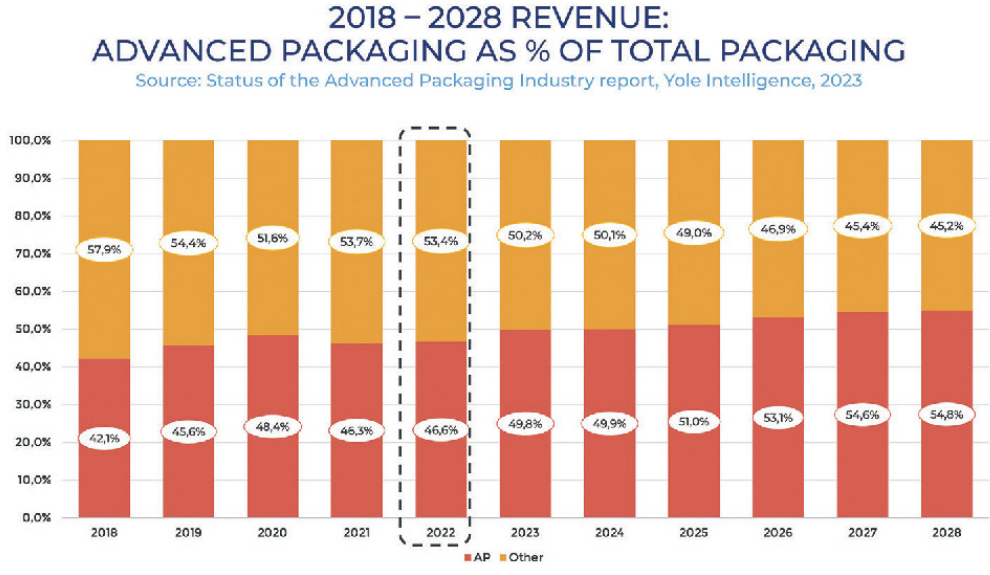


Figure 3. Advanced packaging market share (%) evolution from 2018 to 2028 in revenues.

Continued from page 39

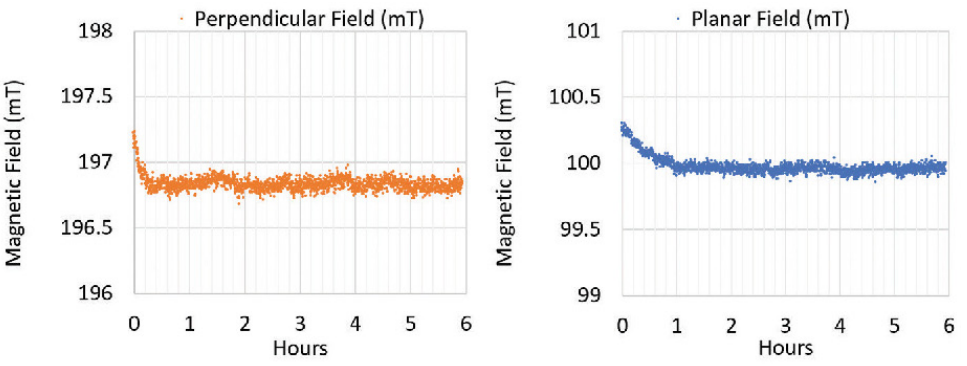


Figure 4. Illustration of static field stability while applied continuously for 6 hrs.

end-application. **FIGURE 4** illustrates the application of static out of plane and static in-plane field component at 200mT. The field is shown here to be applied continuously for 6 hours. After the first stabilization period, field variation within 0.1mT is shown. Stress with vectorial field at controlled and variable angle can be also executed like on a rotating field pattern illustrated in **FIGURE 5**. Dynamic out of plane rotating field with constant norm of amplitude at above 100mT is plotted to simulate a vector field coming from any direction in a plane that is perpendicular to the chip surface. The system can expose the MTJ by controlling the field strength, angle of vector and duration of exposition simultaneously with BER tests.

Summary

STT-MRAM technology is revolutionizing traditional charged based memory and its large deployment is now initiated. The driving applications for its deployments are today wearable, IoT and automotive. We illustrated empirically that in a car, the parasitic magnetic fields measured are very low compared to the switching field of the free layer. If the source of field is a permanent magnet, then the distance to the magnet must be at a controlled distance as the magnetic field decrease very quickly with it. MRAM manufacturers specify the magnetic immunity at a bit error rate of less than 1ppm at magnetic fields amplitude levels which are 20 to 100 times larger than all the measurements performed here. Nevertheless, for automotive applications, it is mandatory for safety of the system that data are safely stored and available in the memory [18]. To control STT-MRAM magnetic immunity during the manufacturing process, the wafers are stressed by a vectorial magnetic field of controlled strength and direction to evaluate the Bit Error Rate. The level of resilience to a certain level of magnetic field and specific angles are then to be compared with the context of end application. If the MRAM chip has immunity specifications at below the level of magnetic field present in the end application, several solutions are already available such as the use of magnetic shielding in the chip packaging. System level consideration also enables several possibilities like guarding, by design, a distance to avoid any presence of permanent magnet. These solutions being on top of already implemented Error Correction Code (ECC) algorithms used to correct unwanted switching events within the memory.

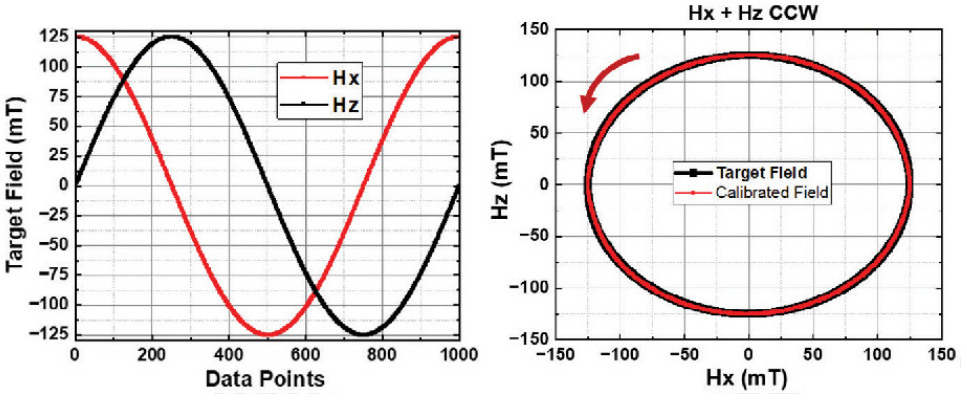


Figure 5. Illustration of rotating field in the XY direction for angular resolved immunity tests.

REFERENCES

1. https://semiengineering.com/embedded-flash-scaling-limits/
2. https://ambiq.com/apollo4-plus/
3. https://ambiq.com/news/xiaomi-and-ambiq-collaborate-to-create-new-redmi-watch-3/
4. C. -H. Chen et al., "Reliability and magnetic immunity of reflow-capable embedded STT-MRAM in 16nm FinFET CMOS process," 2021 Symposium on VLSI Technology, Kyoto, Japan, 2021, pp. 1-2.
5. T. Y. Lee et al., "World-most energy-efficient MRAM technology for non-volatile RAM applications," 2022 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 2022, pp. 10.7.1-10.7.4, doi: 10.1109/IEDM45625.2022.10019430.
6. https://www.nxp.com/company/about-nxp/nxp-and-tsmc-to-deliver-industrys-first-automotive-16-nm-finfet-embedded-mram:NW-NXP-AND-TSMC-DELIVER-FIRST16NM-FINFET-MRAM
7. R Pääkkönen , L Korpinen, LOW FREQUENCY MAGNETIC FIELDS INSIDE CARS, Radiation Protection Dosimetry, Volume 187, Issue 2, December 2019, Pages 268–271, https://doi.org/10.1093/rpd/ncz248
8. https://www.emf-portal.org/en/emf-source/71160
9. https://www.senis.swiss/wp-content/uploads/2022/08/3-Axis-Analog-Magnetic-Transducer-I3A-03C02F-B02T0K5J_rev-01.pdf
10. https://www.supermagnete.fr/eng/faq/How-do-you-calculate-the-magnetic-flux-density
11. https://memxcell.com/mram1.html
12. C. -Y. Wang et al., "Impact of external magnetic field on embedded perpendicular STT-MRAM technology qualified for solder reflow," 2017 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 2017, pp. 21.1.1-21.1.4, doi: 10.1109/IEDM.2017.8268432.
13. Y. Ji et al., "Reliability of Industrial grade Embedded-STT-MRAM," 2020 IEEE International Reliability Physics Symposium (IRPS), Dallas, TX, USA, 2020, pp. 1-3, doi: 10.1109/IRPS45951.2020.9129178.
14. V. B. Naik et al., "STT-MRAM: A Robust Embedded Non-Volatile Memory with Superior Reliability and Immunity to External Magnetic Field and RF Sources," 2021 Symposium on VLSI Technology, Kyoto, Japan, 2021, pp. 1-2.
15. Y. -D. Chih et al., "13.3 A 22nm 32Mb Embedded STT-MRAM with 10ns Read Speed, 1M Cycle Write Endurance, 10 Years Retention at 150°C and High Immunity to Magnetic Field Interference," 2020 IEEE International Solid- State Circuits Conference - (ISSCC), San Francisco, CA, USA, 2020, pp. 222-224, doi: 10.1109/ISSCC19947.2020.9062955.
16. T. Y. Lee et al., "Magnetic Immunity Guideline for Embedded MRAM Reliability to Realize Mass Production," 2020 IEEE International Reliability Physics Symposium (IRPS), Dallas, TX, USA, 2020, pp. 1-4, doi: 10.1109/IRPS45951.2020.9128317.
17. Y. Ji et al., "Reliability of 8Mbit Embedded-STT-MRAM in 28nm FDSOI Technology," 2019 IEEE International Reliability Physics Symposium (IRPS), Monterey, CA, USA, 2019, pp. 1-3, doi: 10.1109/IRPS.2019.8720429.
18. https://www.pcmag.com/news/tesla-asked-to-recall-158000-vehicles-over-flash-memory-failure

Continued from page 36

inventories clearly shows up from the middle of 2022, when parts inventories held by OEMs spiked up to 40% of revenues. Slowly increasing lead times for many parts acted as another catalyst for maintaining higher stock levels.

Interestingly, by Q3 2022, it was clear that orders for equipment were starting to falter, yet OEMs continued to order parts as they rushed to clear their backlogs before customers canceled them. During this period, OEMs also signaled to their suppliers to keep building and shipping parts. When the reality of falling orders and cancellations hit and translated into falling revenues at the beginning of 2023, the OEMs slammed on the brakes. This action came too late, and inventories of parts held by OEMs jumped to an unsustainable 48% of revenues. The expectation is that parts inventories held by OEMs will remain persistently high for the next two quarters until stocks run down and demand for wafer fab equipment recovers.

From the suppliers' perspective, the rapid change in inventories they now hold has come as a shock. Historically, their inventory of finished products would typically be in the 20% to 25% range, but in recent years this has fallen to 15% to 10% of revenues as demand outstripped supply. Now, because OEMs encouraged suppliers to keep producing despite their orders falling, inventories of finished goods held by suppliers have leapt to 20% and are expected to grow further (**FIGURE 1**).

So far, higher inventories are not a severe problem for suppliers as revenues are still high. Q1 2023 revenues, while 23% below peak quarterly revenues, is still the 3rd best quarter ever. Even with revenues expected to fall a further 5% to 10% in Q2, most suppliers will still be profitable. In addition, most suppliers are using this quiet period to fix problems they didn't have time to address during the ramp. These problems include product development, improving manufacturing efficiencies, ramping up new facilities, finding new partners, and training.

The semiconductor capital equipment market is cyclical, so the supply chain is well structured to deal with this scenario. The difference this time around is the market is so much larger than before and, therefore, potentially carries a higher level of risk if the downturn persists. Fortunately, the current outlook is for a recovery later this year, followed by two to three years of single-digit growth before the next downturn occurs in either 2026 or 2027.

One thing is clear, when the market for wafer fab equipment starts to ramp up when memory prices recover, and demand for chips related to generative AI takes off, the supply chain will be more robust and better adapted to respond effectively and efficiently than ever before.

SOURCES:

- Semiconductor Equipment Subsystems Market Monitor, Yole Intelligence
- Semiconductor Test Equipment Market Monitor, Yole Intelligence
- Wafer Fab Equipment Market Monitor, Yole Intelligence

About the author

John West is a Senior Director of the Semiconductor Subsystems and Test Division at Yole Group. He has over 20 years of industry experience and a successful track record in various strategy and consulting projects. John has a Bachelor's degree in Medical Physics from King's College London and an MBA from Cranfield School of Management.

AD INDEX

Advertiser	Pg
Atlas Technologies	13
Busch LLC	7
Edwards Ltd.	3
Electronic Fluorocarbons	9
Elevate Semiconductor	18-19
EOS/ESD Association, Inc.	25
Levitronix Technologies, Inc.	C3
Middlesex General Industries, Inc.	11
Nikon Precision, Inc.	C2
Nordson Test & Inspection	5, C4
Pfeiffer Vacuum	55
SEMI	23
Stratus Automation SDN BHD	17
Thermcraft, Inc.	15

SEMICONDUCTOR DIGEST
NEWS AND INDUSTRY TRENDS

OFFICES

Gold Flag Media, 58 Summer Street, Andover, MA 01810

ADVERTISING

North America, ROW

Kerry Hoffman
Tel: 978-580-4205
kerryh@semiconductordigest.com

Germany, Austria, Switzerland & E. Europe

Holger Gerisch
Tel: +49 0 8856 8020228
holger.gerisch@gmx.de

For advertising information, please visit www.semiconductordigest.com/advertise.

Continued from page 52

growth rate of approximately 6%. This growth can be attributed to the technological trends mentioned above and the critical investments in the AP landscape.

Among various AP platforms, the highest revenue CAGR rates are expected from Embedded Die (ED), 2.5D/3D stacking, and flip-chip, with anticipated growth rates of 30%, 18.7%, and 9%, respectively. These technologies are witnessing increased market penetration, with 3D stacking finding applications in AI, High-Performance Computing (HPC), data centers, Complementary Metal-Oxide-Semiconductor Image Sensors (CIS), and 3D NAND. At the same time, ED is being adopted in the automotive, mobile, and base station sectors.

The mobile and consumer market constituted most of the total revenue generated by advanced packaging in 2022. The fastest-growing segment in terms of revenue is expected to be telecom and infrastructure, with an estimated growth rate of approximately 17%, leading to an increased market share from 20% in 2022 to 27% by 2028. Furthermore, the automotive and transportation segment is forecasted to grow at a CAGR of 10% from 2022 to 2028, reaching around **$7 billion** in revenue by 2028.

In 2022, flip-chip technology contributed approximately 51% of the revenue among various AP technologies. However, its market share is projected to decrease to around 47% by 2028, while the 2.5D/3D stacking percentage is expected to increase from approximately 21% in 2022 to 33% by 2028. 2.5D/3D stacking is forecasted to continue its impressive growth, with a CAGR of roughly 30%, driven by its adoption in AI, HPC, 3D NAND, 2.5D interposer-based die partitioning, heterogeneous integration, 3D System-on-Chip (SoC), and stacked CIS. The fan-out packaging market, already small, is expected to decrease its market share from 4% in 2022 to 3% by 2028. Conversely, embedded die is projected to grow at a CAGR of 30% over the next five years, driven by demand from sectors such as telecom and infrastructure, automotive, and medical.

In conclusion, the advanced semiconductor packaging industry is witnessing significant advancements and resilient growth. Advanced packaging technologies are well-positioned to address the evolving demands of high-performance computing, AI, and 5G. Despite an economic slowdown in the semiconductor industry, advanced packaging achieved remarkable revenue growth of 10% in 2022. It is projected to surpass traditional packaging in revenue by 2025. Taiwanese companies, led by ASE, are dominating the market, but semiconductor giants like TSMC, Intel, and Samsung are actively participating and driving innovation. Continuous technological advancements in processes, materials, and equipment are essential for the future success of advanced packaging, enabling heterogeneous integration and meeting the demands of next-generation hardware. The industry's growth and resilience indicate a promising future characterized by ongoing collaboration and innovation among key players.

This article has been written in collaboration with Emilie Jolivet, Director, Semiconductor, Memory & Computing at Yole Intelligence. Yole Intelligence is a Yole Group company.

SEMICONDUCTOR DIGEST PRODUCT SHOWCASE

Magnetically Levitated Turbopumps

The Pfeiffer Vacuum ATH 2804 M/MT and ATH 3204 M/MT magnetically levitated turbopumps are designed for harsh semiconductor applications. They offer a gas throughput of over 5,000 sccm of nitrogen for non-heated applications and up to 1,500 sccm of argon at 65°C in corrosive applications. To prevent by-product deposition, the MT model can be heated and regulated up to 85°C.

PFEIFFER VACUUM

www.pfeiffer-vacuum.com

Indium Plating Challenges: Unlocked

JOHN GHEKIERE, Vice President, Product & Technology, ClassOne Technology

INDIUM IS THE GREAT APPEASER OF interconnect materials. However, it also tends to be a bit of an attention seeker. One day it's a hot topic and everyone's interested. The next, it's suddenly passé, and you get eye rolls when you mention it. Today, indium's time may finally be here.

As a solder or even a seal, indium presents an intriguing case as an enabling material. Pillar-based methods of packaging devices, which we're still calling advanced, have become commonplace as a means of stacking die for increased performance. But now, the industry is also experiencing a persistent, explosive expansion of device types and use cases. And the harsh truth is that if you make two perfect semiconductor devices with 100% yield each, and then you connect them through pillars that don't hold up, your amazing chips become useless.

Chip packages are showing up everywhere now, including the most extreme conditions into which we can inject, install or launch them. Take satellites. That expensive cargo, folded neatly inside the nose cone of a rocket, launches from the pad at a temperature of maybe 24°C. In orbit, it neatly and cleverly unfolds itself in an environment of around -270°C. As it zips from dark side to sunny side of the Earth, radiation from the sun raises its temperature back up by some 200+ degrees.

Common solders readily crack in these conditions, forcing aerospace engineers to burden an otherwise trim piece of technology with a lot of bulk and complexity (and mass!) for heat deflection and abatement.

Enter indium, the appeaser — it gets along with everyone in almost any situation. Indium, magically, joins the scene with a malleability that fits almost in the middle of what was historically the precious sweet spot of lead-based solders. Like a gluten-free bread that actually tastes ok, indium provides quite favorable mechanical characteristics like a conventional solder. Amazingly, it retains these at near-absolute zero and holds up over time through the kind of heat cycling that low-earth orbit imposes.

But… there has to be a downside, or everyone would be using indium. As I hinted at the start, there's been some drama.

In the mid 2010s, rumors arose of an impending shortage of indium. Indium is included on the U.S. government's list of Technology-Critical Elements (TCE), and in recent years, its refinement infrastructure has caught up fully to demand. While still somewhat rare, indium is more readily available than gold, platinum or iridium.

Deposition at small feature sizes becomes the next hurdle. Resistive thermal evaporation — historically, the leading method — bears a number of brutal shortcomings for volume manufacture, including overall cost and flexibility. These challenges worsen as features are scaled, even modestly. On the other hand, electroplating of indium is highly flexible and cost-manageable.

Indium plates beautifully. But here, too, there have been complications (spoiler alert: we've figured them out). The two key issues: very low plating efficiency, which leads to excessive gas evolution, thus void defects, etc.; and enormous as-deposited grains. Plating efficiency and gas evolution are related and, in large part, tied to chemistry formulation. However, even after formulation improvements by several chemical manufacturers, gas evolution was still not completely eliminated. Indium incurs a galvanic reaction in contact with a host of other metals when submerged. So, getting the reactor and anode design right is critical — not a straightforward task since indium wets certain other solids, and will fold itself around substructure.

Once all the design aspects are sorted (which we've done), you still have the grain-size issue. It is not uncommon to pursue a packaging scheme deployed around pillars 20µm in diameter or even smaller. Native grain size of electrodeposited indium tends to be on the order of 30-40µm! That means a single grain is, on average, larger than the width of what it fills. Not great.

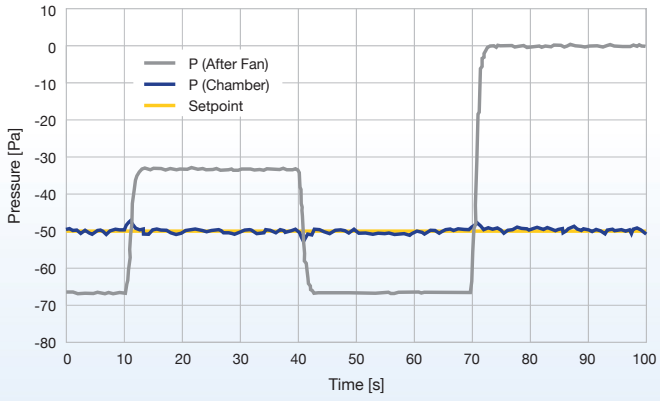
In the end, for manufacturers interested in exploring indium's benefits, the key is to start with a technology partner that understands this material well and can serve as a center point for pulling it all together: chemistry, wafer and reactor technology. A partner used to the indium drama. ClassOne Technology sees ourselves as just that kind of partner. SD

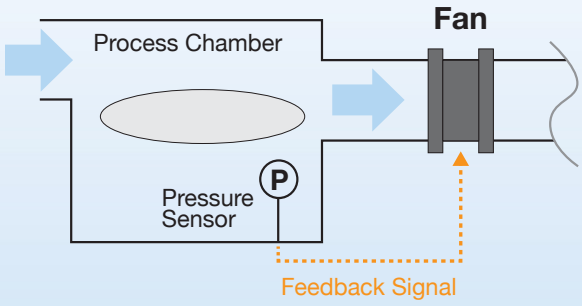
LEVITRONIX®

Better Fans for Better Yield

DON´T LET YOUR EXHAUST CONTROL YOUR PROCESS!

Keep Your Chamber Pressure constant with the Levitronix® MagLev Fan!
Placed on the exhaust line, the fan compensates for pressure fluctuations due to an inconsistent or insufficient central exhaust.





LEVITRONIX® MAG*LEV* FAN

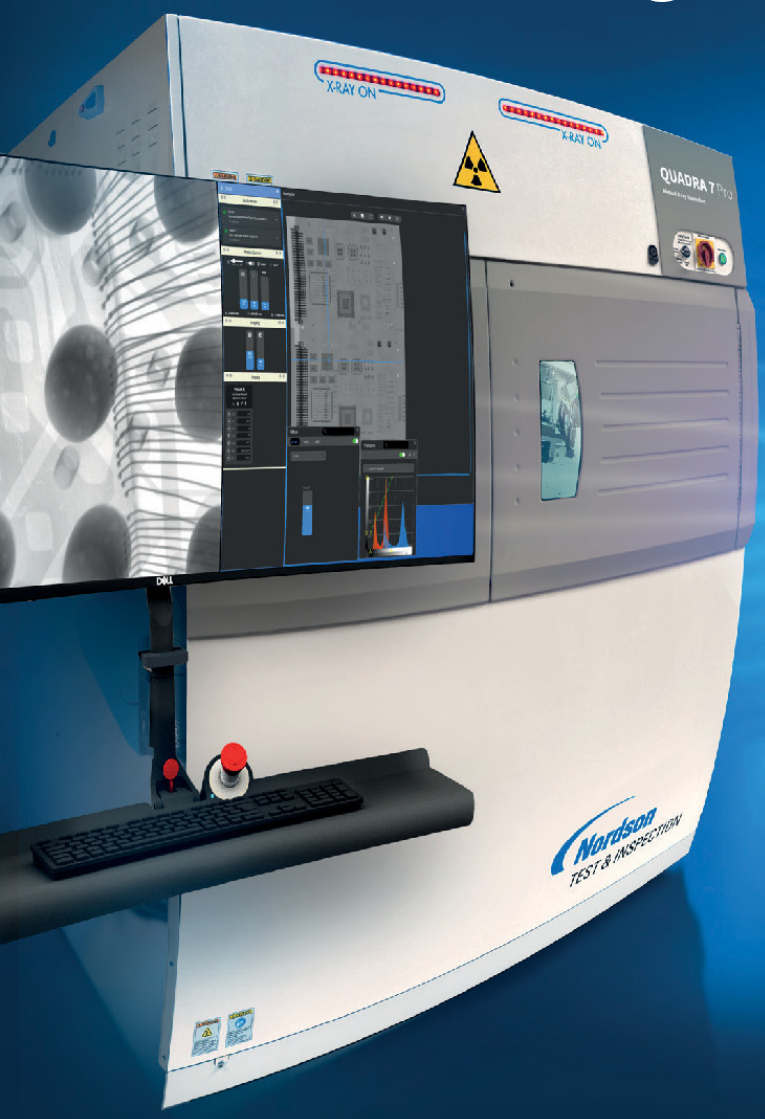
- ATEX/IECEX certified
- High flow capabilities
- Chemically resistant
- Low maintenance
- Magnetically levitated rotor

www.levitronix.com

Making the Invisible Visible

QUADRA® Pro

Manual X-ray Inspection



Fulfill your sensitive semiconductor inspection with even higher resolution images

Proprietary, high performance dual mode X-ray tube and a new advanced detector speed revealing the tiniest details

Revalution™ control software features simple intuitive menus and controls, even with multiple complex images

nordson.com/TestInspect

Nordson
TEST & INSPECTION