SILICON SEMICONDUCTOR

ISSUE VIII 2026

AN ANGEL BUSINESS COMMUNICATIONS PUBLICATION

SILICONSEMICONDUCTOR.NET



DEMANDS FOR HIGHER PRODUCTIVITY, EQUIPMENT RELIABILITY, AND SUSTAINABILITY ARE DRIVING A GLOBAL SHIFT TO PREDICTIVE MAINTENANCE

EDWARDS

SCREEN

Driving Innovation Aimed at Net Zero with Our Partners

Go Green toward Net Zero

Provision of eco-friendly products and services

The SCREEN Group is working to expand sales of products that are certified as Green Products according to its independent evaluation standards for excellent environmental performance.



Single Wafer Cleaner
SU-3200



Spin Scrubber
SS-3200



Coat/Develop Track
DT-3000

SCREEN Semiconductor Solutions Co., Ltd.

SEMICON® EUROPA Nov 10–13, 2026, SCREEN Booth #B1700

VIEWPOINT

BY PHIL ALSOP, EDITOR



Semiconductor strategy becomes the defining competitive advantage

If there is a common thread running through this month's semiconductor headlines, it is that the industry has entered a new phase where scale, collaboration and strategic national ambition matter just as much as technology breakthroughs themselves.

From Singapore's launch of the new SG Semiconductor national brand to multibillion-dollar investments in manufacturing, R&D and AI infrastructure across the United States and Asia, the sector is increasingly being shaped by ecosystems rather than individual companies.

Singapore's initiative is perhaps the clearest example of this trend. While the city-state has been a major semiconductor hub for decades, accounting for a significant share of global chip output and semiconductor equipment production, SG Semiconductor represents something more strategic than a branding exercise. It creates a single identity that brings together research, manufacturing, talent development and commercialisation under one umbrella. At a time when countries worldwide are competing to attract investment and critical technologies, clarity of purpose and visibility are becoming powerful competitive assets.

The announcement was accompanied by a series of industry partnerships spanning advanced packaging, silicon photonics and sensing technologies. That emphasis on collaboration is significant. Increasingly, semiconductor innovation depends on the ability to move technologies seamlessly from laboratory research to commercial-scale manufacturing. Nations and regions that can provide this complete pathway will be best placed to capture future investment.

The same theme emerges elsewhere. Applied Materials' new $5 billion EPIC Centre in Silicon Valley is designed around the idea that no single organisation can solve the challenges of next-generation semiconductor development alone. By bringing chipmakers, equipment suppliers, universities and systems companies together under one roof, the facility reflects a growing recognition that future advances in AI-era computing will require unprecedented levels of cross-industry cooperation. Similarly, the decision by TSMC and ASML to spearhead development of large-format 12-inch photomasks for High-NA EUV illustrates how the industry is beginning to prepare for the manufacturing requirements of the next decade. The project may not deliver production systems until the early 2030s, but it demonstrates the extraordinary long-term planning that now underpins leading-edge semiconductor manufacturing.

Meanwhile, recent data from Dell'Oro Group highlights the demand-side forces driving these investments. A 182% surge in data centre semiconductor and component revenues underlines the continuing impact of AI infrastructure spending. Notably, growth is no longer confined to accelerators alone. Memory, storage and general-purpose server technologies are also benefiting, reflecting the reality that AI requires a complete infrastructure stack.

Perhaps most revealing is Cambridge Consultants' finding that nearly six in ten executives view radical innovation as their most exciting growth opportunity. However, the report also highlights a persistent execution gap. Organisations increasingly understand the importance of breakthrough innovation, yet many still struggle with the cultural, financial and organisational structures needed to bring advanced technologies to market.

That challenge helps explain the growing focus on ecosystem building. Whether through Singapore's coordinated national approach, Applied's collaborative R&D model or the ambitious U.S.-Japan partnership that will see GlobalFoundries operate major new fabs in New York, the industry is recognising that innovation flourishes when talent, capital, manufacturing and research are aligned.

The semiconductor sector has always depended on collaboration, but today that collaboration is becoming more structured, more strategic and more global.

As AI, advanced computing and digital infrastructure continue to expand, the winners may not simply be those with the best technology, but those that build the strongest ecosystems around it. In the semiconductor industry's next chapter, strategy and cooperation are becoming as important as the silicon itself

For your free weekly e-news roundup go to: **siliconsemiconductor.net**

Contents

COVER STORY

Demands for higher productivity, equipment reliability, and sustainability are driving a global shift to predictive maintenance

With rising chip demand and advanced device architecture, fabs face unprecedented complexity in production. The need for higher productivity, increased equipment reliability, and sustainable solutions is driving a global shift toward Smart Manufacturing.



16 Available is not in control

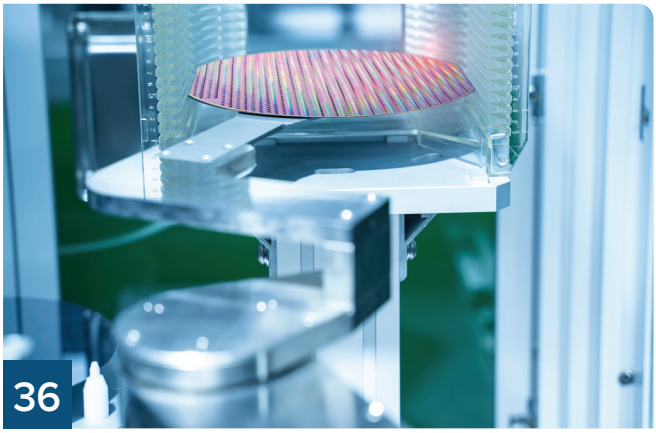
Balancing output, quality, and risk in high-volume fabs.

18 Precision timing: unlocking efficiency and advancing semiconductor sustainability

As AI infrastructure, data centres and edge computing continue their rapid scale-up, the systems that win on efficiency will be those in which timing shapes architecture from the outset of design.

20 CEA-Leti advances alternative computing architectures

Two recent developments from CEA-Leti highlight the breadth of innovation occurring beyond conventional CMOS scaling.



22 Before the first weld

What fab gas-line teams should write down when selecting enclosed orbital welding equipment.

24 How California's capital region is emerging as a major chip hub

The Greater Sacramento Economic Council, outlines the many factors which are contributing to the rise of Sacramento as a globally significant semiconductor hub – not least its proximity to Silicon Valley.

30 Accelerating the next era of custom AI silicon

TYLsemi is the first chiplet platform company to deliver a full portfolio across IO, power delivery, and memory paired with a chiplet-based custom silicon design, integration and supply-chain ownership.

36 The hidden impact of lubricant selection on contamination control and yield in semiconductor fabs

A technical look at how lubricant chemistry, outgassing behavior, and particulate control influence semiconductor tool performance.

40 AI-designed silicon: Redefining the route from workload to accelerator

The development of specialised AI hardware has traditionally been a lengthy, sequential process.



44 Rethinking CMP for advanced packaging

How flexible, data-driven and resource-efficient platforms can support chiplets, 3D integration and panel-level packaging.

48 How to improve active photonics alignment performance

Latest advancements in algorithms and motion systems.

56 Turning water into a competitive advantage for the semiconductor industry

The semiconductor industry is entering a period of growth that would have looked implausible just a few years ago. The industry is on track to reach $1 trillion in annual revenue by 2030, and demand for AI-capable chips alone has grown by roughly 50% in the past three years.

58 Preventing downtime starts before a semiconductor fab starts operating

The semiconductor industry's race to add capacity will continue as AI and other advanced technologies increase chip demand.

NEWS

06 Singapore launches SG Semiconductor

08 Nearly six in ten CXOs are excited by the opportunity of radical innovation as a growth strategy

09 Data centre IT semiconductor and component revenue surges 182%

10 Applied Materials unveils $5 Billion EPIC Center to accelerate semiconductor innovation for the AI Era



60 SCREEN's holistic path to sustainable semiconductor manufacturing

As demand for advanced chips accelerates, manufacturers must simultaneously increase production capacity while dramatically reducing their environmental footprint. Water consumption, wastewater generation, resource efficiency, and carbon reduction have become strategic priorities across the semiconductor value chain.

Publisher Jackie Cannon jackie.cannon@angelbc.com +44 (0)1923 690205
Editor Phil Alsop phil.alsop@angelbc.com
Editorial Assistant Shehzad Munshi shehzad.munshi@angelbc.com +44 (0)1923 690215
Director of Logistics Sharon Cowley sharon.cowley@angelbc.com +44 (0)1923 690200
Design & Production Manager Mitch Gaynor mitch.gaynor@angelbc.com +44 (0)1923 690214
Graphic Design & Multimedia Assistant Harvey Watkins harvey.watkins@angelbc.com

Chief Executive Officer Sukhi Bhadal sukhi.bhadal@angelbc.com +44 (0)2476 718970
Chief Technical Officer Scott Adams scott.adams@angelbc.com +44 (0)2476 718970
Directors Jackie Cannon, Sharon Cowley

Published by Angel Business Communications Ltd, 6 Bow Court, Fletchworth Gate, Burnsall Road, Coventry CV5 6SP, UK.
T: +44 (0)2476 718 970 E: info@angelbc.com

Silicon Semiconductor is published ten times (6x print and 4x video issues) a year on a controlled circulation basis. Non-qualifying individuals can subscribe at: £50.00 per annum (UK) , €60.00 per annum (Europe), $90 per annum (air mail)(USA). Cover price £4.50. All information herein is believed to be correct at time of going to press. The publisher does not accept responsibility for any errors and omissions. The views expressed in this publication are not necessarily those of the publisher. Every effort has been made to obtain copyright permission for the material contained in this publication. Angel Business Communications Ltd will be happy to acknowledge any copyright oversights in a subsequent issue of the publication. Angel Business Communications Ltd © Copyright 2026. All rights reserved. Contents may not be reproduced in whole or part without the written consent of the publishers. The paper used within this magazine is produced by chain of custody certified manufacturers, guaranteeing sustainable sourcing. US mailing information: Silicon Semiconductor, ISSN 1096-598X, is published 6 times a year by Angel Business Communications Ltd, Unit 6, Bow Court, Fletchworth Gate, Burnsall Rd, Coventry CV5 6SP. UK. Airfreight and mailing in the USA by agent named World Container INC 150-15, 183rd St, Jamaica, NY 11413, USA. Periodicals Postage Paid at Brooklyn, NY 11256. POSTMASTER: Send address changes to Silicon Semiconductor, Air Business Ltd, c/o World Container INC 150-15, 183rd St, Jamaica, NY 11413, USA.

We strive for accuracy in all we publish; readers and contributors are encouraged to contact us if they recognise an error or omission. Once a magazine edition is published [online, in print or both], we do not update previously published articles to align old company names, branding, marketing efforts, taglines, mission statements or other promotional verbiage, images, or logos to newly created or updated names, images, typographic renderings, logos (or similar) when such references/images were accurately stated, rendered or displayed at the time of the original publication. When companies change their names or the images/text used to represent the company, we invite organizations to provide Angel Business Communications with a news release detailing their new business objectives and/or other changes that could impact how customers/prospects might recognise the company, contact the organisation, or engage with them for future commercial enterprise.
Printed by: The Manson Group. ISSN 2050-7798 (Print) ISSN 2050-7801 (Online) © Copyright 2026.

Singapore launches SG Semiconductor

National brand provides a clearer gateway to Singapore's capabilities, partners and talent, alongside new and expanded industry collaborations.

MINISTER for Trade and Industry (Energy and Industry) Dr Tan See Leng has announced SG Semiconductor, a national identity for Singapore's semiconductor sector, at Innovate Together 2026. Organised by the Agency for Science, Technology and Research (A*STAR), the annual forum connects public research and industry to advance semiconductor technologies.

Jointly developed by A*STAR and the Singapore Economic Development Board (EDB), SG Semiconductor will raise the global visibility of Singapore's sector-wide strengths and make it easier for companies and talent to understand what the country offers and engage its ecosystem. The branding effort will focus on public sector R&D as its first course of action.

Singapore has built a substantial semiconductor industry over nearly six decades, spanning chip design, wafer fabrication, assembly and testing, equipment and materials development and manufacturing, and research and development (R&D). It now accounts for one in ten global chip output and one-fifth of global semiconductor manufacturing equipment production. As applications become more sophisticated and place greater demands on semiconductor performance, Singapore is building on this foundation through closer research-industry collaboration and talent development, supported by the latest S$800 million investment from 2026 to 2030 through the RIE Flagship in Semiconductors.

Trusted location for semiconductor innovation and manufacturing

SG Semiconductor brings together Singapore's capabilities across R&D, advanced manufacturing, infrastructure, industry partnerships and talent. Its work spans seven technology areas: advanced packaging, silicon photonics, power electronics, radio frequency gallium nitride, piezoelectric micro-electro-mechanical systems (piezoMEMS), flat optics and integrated circuit design. Across these areas, Singapore is developing capabilities to address growing needs in artificial intelligence and advanced computing, high-speed and energy-efficient data communications, electrification, sensing and intelligent systems. By connecting research expertise with industry-grade infrastructure and manufacturing capabilities of industry partners, Singapore can develop, validate and scale new semiconductor technologies for global markets.

"Singapore's strength in semiconductor innovation has been built through decades of sustained investment and close partnership across public research, universities and industry. SG Semiconductor brings this collective endeavour under a national identity. It signals our commitment to build frontier capabilities, translate them into industry outcomes and anchor high value R&D and manufacturing in Singapore for the long term," said Mr Beh Kian Teik, Chief Executive, A*STAR - Agency for Science, Technology and Research (A*STAR).

"As semiconductor technologies become more complex, companies are increasingly seeking trusted locations where innovation, manufacturing and talent come together. SG Semiconductor provides a gateway to Singapore's ecosystem, enabling companies to develop, scale and



commercialise next-generation technologies from Singapore." said Mr Jermaine Loy, Managing Director, EDB.

Accelerating innovation through new industry partnerships

Demonstrating SG Semiconductor in action, Innovate Together 2026 also saw the announcement of new and expanded partnerships spanning several key semiconductor technology areas. These collaborations connect research expertise with industry-grade platforms and manufacturing capabilities to develop, validate and scale emerging technologies.

In advanced packaging, Applied Materials and A*STAR will embark on Phase 4 of their longstanding collaboration, significantly expanding the joint laboratory's infrastructure, equipment and headcount to develop technologies for AI and advanced computing. A*STAR and KLA will establish a new process control collaboration framework to explore ways to improve manufacturing reliability and yield, while A*STAR and STATS ChipPAC will collaborate in copackaged optics to create a pathway towards high-volume production.

Together, these collaborations strengthen Singapore's position as a hub for advanced packaging innovation and help accelerate the translation of next-generation semiconductor technologies into manufacturing. In silicon photonics, A*STAR and GlobalFoundries will deepen their R&D collaboration to develop next-generation technologies on 300 mm wafers in Singapore, supporting higher-volume manufacturing for high-performance computing and data communications. This will help bridge the gap between laboratory innovation and commercial-scale production, supporting growing global demand for high-speed, energy-efficient connectivity solutions.

In MEMS sensing, Tacta Systems and A*STAR will develop intelligent sensing technologies for robotics, tapping A*STAR's Lab-in-Fab platform, which brings together research and manufacturing capabilities for industry-relevant development, prototyping and validation. Tacta Systems has also opened its Singapore operations, strengthening its technology development activities here.

The wider ecosystem represented by SG Semiconductor includes a new research collaboration between National Centre for Advanced Integrated Photonics (NCAIP), a national centre hosted at Nanyang Technological University (NTU) and Battery Age Minerals. The partners will explore germanium-based devices for data communications by combining NTU's research expertise with Battery Age Minerals' access to raw germanium, creating a potential pathway from upstream resources to higher-value semiconductor applications. The collaboration highlights how Singapore's ecosystem brings together partners across the global value chain to explore new opportunities and create future sources of semiconductor innovation.

"These partnerships demonstrate Singapore's ability to bring together world-class research, industry-grade platforms and advanced manufacturing capabilities to move emerging technologies from concept to commercialisation. SG Semiconductor will strengthen this advantage by giving companies a clearer path to translate innovation into globally competitive products and scale them for international markets," said Professor Yeo Yee Chia, Deputy Chief Executive (Innovation & Enterprise), A*STAR and Executive Vice President of the RIE2030 Flagship in Semiconductors.

Convening semiconductor leaders at Innovate Together

Now in its second year, Innovate Together has established itself as Singapore's annual forum connecting public research and industry to advance semiconductor technologies. The event brings together local and international leaders from industry, academia and research for executive perspectives, panel discussions, technical talks and technology showcases. Senior leaders from Applied Materials, GlobalFoundries, KLA and STMicroelectronics discussed emerging industry trends and the pathways needed to bring new technologies to market.

Held as part of the SSIA Summit, the event also showcased Singapore's advances across advanced packaging, silicon photonics, power electronics, radio frequency gallium nitride (RF GaN), piezoMEMS and flat optics, and created opportunities for participants to explore new collaborations.

Dedicated webinars for the silicon semiconductor industry

Using our 30+ years' experience in B2B vertical technical markets, and as the publisher of Silicon Semiconductor, we offer effective webinars, ZOOM interview and virtual events. We help you get your message, to your desired audience, by marketing to over 55,000 silicon semiconductor professionals.

Find out how Angel Webinar can help you organise and run a webinar within the silicon semiconductor industry.

Reach Educate Influence
- Brand Awareness
- Lead Generation
- Thought Leadership



SS SILICON SEMICONDUCTOR

Contact: Jackie Cannon
jackie.cannon@angelbc.com
+44 (0)1923 690205

Nearly six in ten CXOs are excited by the opportunity of radical innovation as a growth strategy

Radical innovation is becoming a clear strategic necessity for ambitious global business leaders to address toughening customer demands, geopolitical instability and shareholder desire for continued double-digit growth.

THERE is a significant disconnect between deep tech intent and execution, with organizational, cultural, and external factors proving to be greater barriers to innovation than technology itself.

This is according to a new global Deep Tech Index by Cambridge Consultants (CC), the deep tech powerhouse of Capgemini, which explores what it takes to achieve radical innovation, and some of the most critical barriers faced by C-suite executives today.

Based on a survey of 750 global CXOs and 25 in-depth executive interviews, the report reveals that nearly six in ten (59%) CXOs consider radical innovation the most exciting type of innovation. Most (81%) said that defensible technological IP is important in capturing long-term value for their businesses.

Interviewees stated that business leaders are increasingly turning to radical innovation as both a primary defensive and offensive strategy to survive intensifying market demands, volatile competitive landscapes, and accelerating technological shifts. Despite this, the report finds a challenging disconnect between intent and execution that needs to be addressed urgently.

Decisive shift of focus from activity to outcomes

Based on insight from the in-depth executive interviews, CXOs are struggling to deliver impact from innovation and are actively looking to establish specific organizational conditions that move unproven, early-stage ideas cleanly from concept to a defined, real-world commercial outcome. Four clear indicators of



the shift towards radical innovation emerged.

Innovation is a core survival strategy: CXOs recognize radical innovation as an effective vehicle to achieve the double or triple revenue growth targets often demanded by shareholders. There is a shift from “activity” to P&L-based outcomes: Funding exploratory innovation without clear outcomes is coming to an end.

Today, business cases must clearly demonstrate near-term profitability and market readiness.

Shared-risk partnership models: Fee-for-service models are giving way to co-investment partnerships where external technology providers and enterprises are both accountable for outcomes but also share in success.

Deep customer validation: Radical technology programs are strictly gated, requiring early, rigorous proof of end-user demand before capital commitment.

“It’s exciting to see that business leaders increasingly recognize the value of a deep tech strategy and what is really compelling is the acknowledgement that radical innovation can address real commercial need,” said Richard Traherne, CEO, Cambridge Consultants, part of Capgemini. “Crucially, though, organizations are not losing the global innovation race because competitors possess better ideas, but because competitors are superior at funding, protecting, and commercializing them. The gap isn’t technical, it’s structural.”

Bridging the intent-to-execution gap

The survey found that close to one in six business leaders are committing significant budgets to radical innovation, with a cross-industry average of 16%* reporting that their organizations allocate 70% or more of their innovation budgets to radical innovation. However, this allocation varies widely by industry.

For instance, just 3% of leaders from consumer companies claimed to allocate more than 70% of their innovation to radical bets, whereas 36% of semiconductor leaders claimed to do so. Respondents from industrials (22%) and healthcare (18%) were more likely to invest in radical innovation, while those from materials (14%), communication services (13%) and energy (12%) were behind. No-one from utilities companies reported allocating more than 70%.

In interviews, CXOs stated that corporate boardrooms, restrictive funding structures, and risk-averse cultures can actively smother radical breakthroughs well before they reach the market. Short-termism, approval bottlenecks, consensus culture and talent scarcity can all impact radical innovation initiatives.

Data centre IT semiconductor and component revenue surges 182%

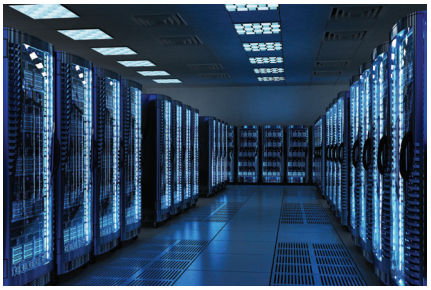
According to a recently published report by Dell'Oro Group, the worldwide Data Center IT Semiconductor and Component revenue for servers and storage systems surged 182 percent in 2Q 2026.

DRAM and storage drives were the strongest growth drivers, while continued AI infrastructure investment fueled broad-based demand across accelerators and complementary server and storage components.

"Data center component growth accelerated sharply in the second quarter, driven by higher memory and storage prices and continued growth in server demand. NVIDIA's Blackwell platform and custom accelerators from Google and Amazon continued to gain momentum," said Baron Fung, Vice President of Research at Dell'Oro Group. "Meanwhile, agentic AI is emerging as a meaningful growth driver for general-purpose servers, broadening component demand beyond accelerated computing. Looking ahead, we expect component revenue growth to maintain momentum through 2026, supported by elevated component prices and strong AI infrastructure investment. Supply constraints, however, may ultimately limit hyperscaler capex plans."

Additional highlights from the 2Q 2026 Data Center IT Semiconductors and Components Quarterly Report:

DRAM and storage drive revenue grew at a triple-digit pace, as average selling prices per bit more than doubled year-over-year.
Accelerated server deployments drove



strong demand for complementary components, including HBM, CPUs, memory, and high-speed back-end NICs.
General-purpose server demand remained strong across cloud service providers, supported by continued infrastructure expansion and emerging agentic AI workloads.

TSMC and ASML pioneer large-format photomasks for High-NA EUV

THE initiative targets establishing a 12-inch mask pilot line by 2031, supporting full lithography system readiness for advanced node production by 2033.

ASML Holding N.V. and Taiwan Semiconductor Manufacturing Company have announced a collaborative initiative to lead the semiconductor industry's transition to a larger-format Extreme Ultraviolet (EUV) lithography photomask.

While High NA EUV will be adopted for production using the current, 6-inch masks, the transition to larger, 12-inch masks is expected to further increase fab productivity, lower chipmaking costs and remove stitching constraints, benefiting the entire semiconductor industry. It will enable advanced chip manufacturers to fully leverage the advantages of High NA EUV, making future generations of leading-edge chips more cost-effective.

"We expect the adoption of High NA EUV to increase progressively along the device scaling roadmap, first using current 6-inch masks and then further supported by 12-inch masks, which enable greater scanner productivity and allow the industry to meet the demand for smaller, faster and more energy-efficient chips." said Christophe Fouquet, president & CEO, ASML. "We are pleased by the strong initial support of semiconductor manufacturers, mask suppliers and partners for this initiative."

"We have always believed that when the industry works together to solve complex problems, we unlock possibilities that no single company could achieve alone," said TSMC Chairman and CEO Dr. C.C. Wei. "By bringing together expertise from across the industry's value chain, we hope to keep providing the benefits of cutting-edge technology through continuous innovation that lowers barriers and puts advanced solutions accessible at scale."

TSMC intends to use ASML's High NA technology in high-volume manufacturing for advanced nodes starting in 2030. As technology nodes advance, TSMC expects the number of layers requiring High NA EUV will rise, driven primarily by the increasingly complex transistor architectures required for AI applications.

On September 7, in advance of the annual SPIE BACUS Conference in Monterey, California, ASML and TSMC established an industry-wide collaborative initiative to drive the transition to 12-inch photomasks. This effort aims to establish a 12-inch mask pilot line by 2031, paving the way for 12-inch High-NA lithography systems to enter advanced node production by 2033.

Underscoring broad industry support, major ecosystem partners and suppliers attended the event and expressed their interest in this transition.

Applied Materials Unveils $5 Billion EPIC Center to Accelerate Semiconductor Innovation for the AI Era

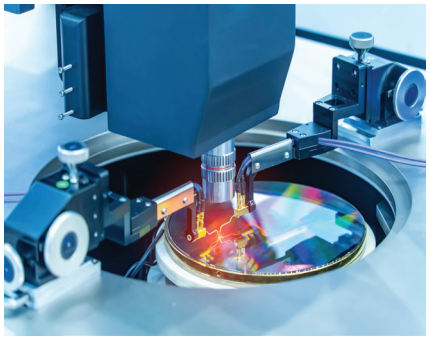
$5 billion investment creates the world's largest semiconductor equipment and process technology R&D facility

APPLIED MATERIALS, Inc., the leader in materials engineering for the semiconductor industry, today unveiled its Equipment and Process Innovation and Commercialization (EPIC) Center, a $5 billion capital investment* to build the world's largest facility dedicated to collaborative semiconductor process technology and manufacturing equipment R&D. Inside the EPIC Center, leading chipmakers, systems companies, equipment partners and university researchers will work side by side with Applied's scientists and engineers to accelerate the pace of chip innovation for the AI era.

"As AI infrastructure scales, the demands of high-performance computing are driving unprecedented complexity across the chipmaking roadmap," said Gary Dickerson, President and CEO of Applied Materials. "To win the race for AI leadership, our industry must rethink and re-engineer how we collaborate to develop and commercialize the next generation of semiconductor technologies. At EPIC, the world's leading chipmakers, system designers, equipment partners and researchers will work side by side with our teams to accelerate the path from breakthrough materials innovation to high-volume manufacturing faster than ever before."

Purpose-Built to accelerate innovation and commercialisation

Unveiled today at Applied's Sunnyvale, California, campus, the EPIC Center represents one of the most complex R&D construction projects in the world. The 600,000-square-foot structure houses nearly four acres of advanced cleanroom space and required more than six million construction hours to complete. Its foundation holds enough concrete to fill 34 Olympic-sized swimming pools, and its steel-reinforced structure holds enough steel



to build two Eiffel Towers. Thirty-four roof trusses cover the cleanrooms, each spanning nearly 300 feet in length and weighing as much as 12 school buses. EPIC is designed to reinvent the R&D model the industry has relied on for decades. In the traditional model, foundational materials research happens in one environment, equipment and process development in another, and process integration and pilot production in a third, with a careful handoff at every phase. The EPIC Center is designed to remove these handoffs, putting the people who invent the materials, build the equipment, and make the chips together in a state-of-the-art facility purpose-built for commercialisation.

EPIC is founded on a simple conviction: the technologies the AI era demands are too complex and too consequential for any one company to invent alone. The EPIC Center opens with a founding roster that spans the AI technology stack. Founding partners TSMC, Intel, SK hynix, Micron and Samsung Electronics are joined by innovation partners Broadcom, Advantest, SCREEN, KIOXIA, and Besi; and by research partners Arizona State University, Rensselaer Polytechnic Institute, Stanford University, UC Berkeley, and UCLA Samueli School of Engineering. Each partner brings its own roadmap challenges and engineers into the facility, working in tight, secure learning cycles alongside Applied teams. Over time, additional companies and organizations are expected to join the EPIC ecosystem.

Leaders from many of the EPIC Center partner companies attended the unveiling event, getting a first glimpse inside the facility. The day started with a summit hosted by Dr. Prabu Raja, President of the Semiconductor Products Group at Applied Materials, where R&D leaders from top semiconductor design, manufacturing and equipment companies discussed key challenges facing growth and progress in the global chip industry.

"Materials innovation has outgrown serial development – at today's atomic-scale dimensions, many materials and materials interfaces behave in entirely new ways, and making them work at industry scale demands coordinated development and integrated solutions," said Dr. Prabu Raja. "EPIC is the only place where materials, process technology and equipment will be innovated in parallel, which increases the success rate of new materials and technologies. Because our customers' own engineers help optimize the technology before it ever reaches their factory, what leaves EPIC arrives at the fab ready to ramp to high performance and high yield sooner."

Applied expects the EPIC operating model to deliver returns on both sides of the partnerships. Chipmakers gain earliest access to next-generation materials and equipment, shorter learning cycles and a faster path to volume production. Applied gains visibility across several generations of emerging technology challenges, informing where the company directs its R&D investments to create value for customers and have its systems and services designed in.

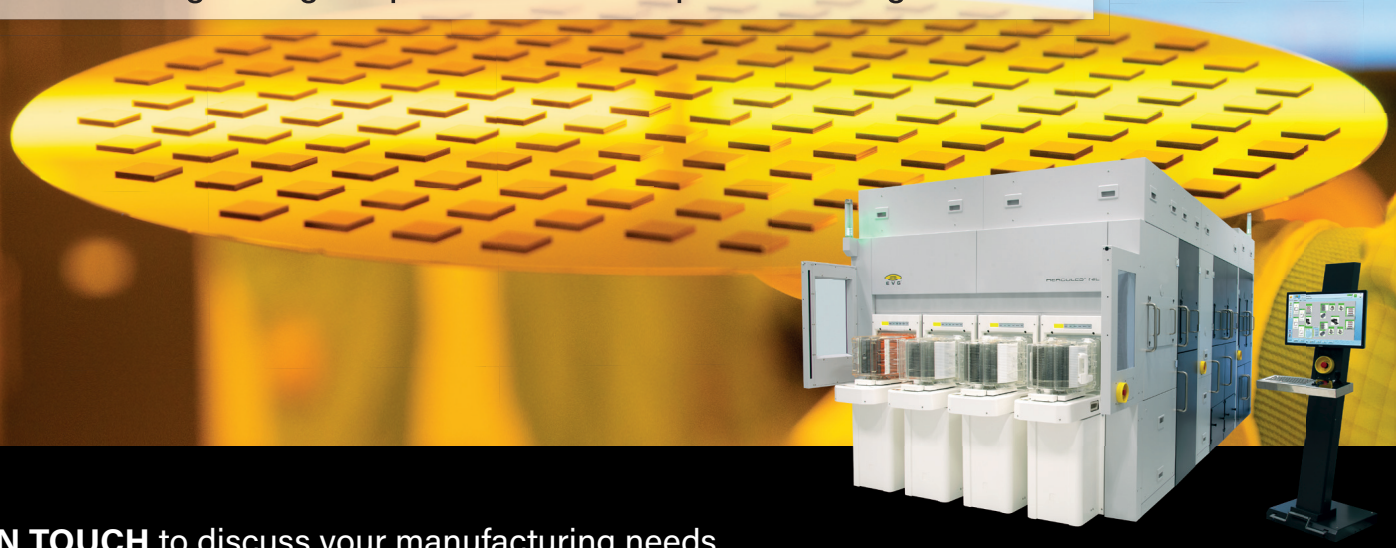
EVG®

invent
innovate
implement

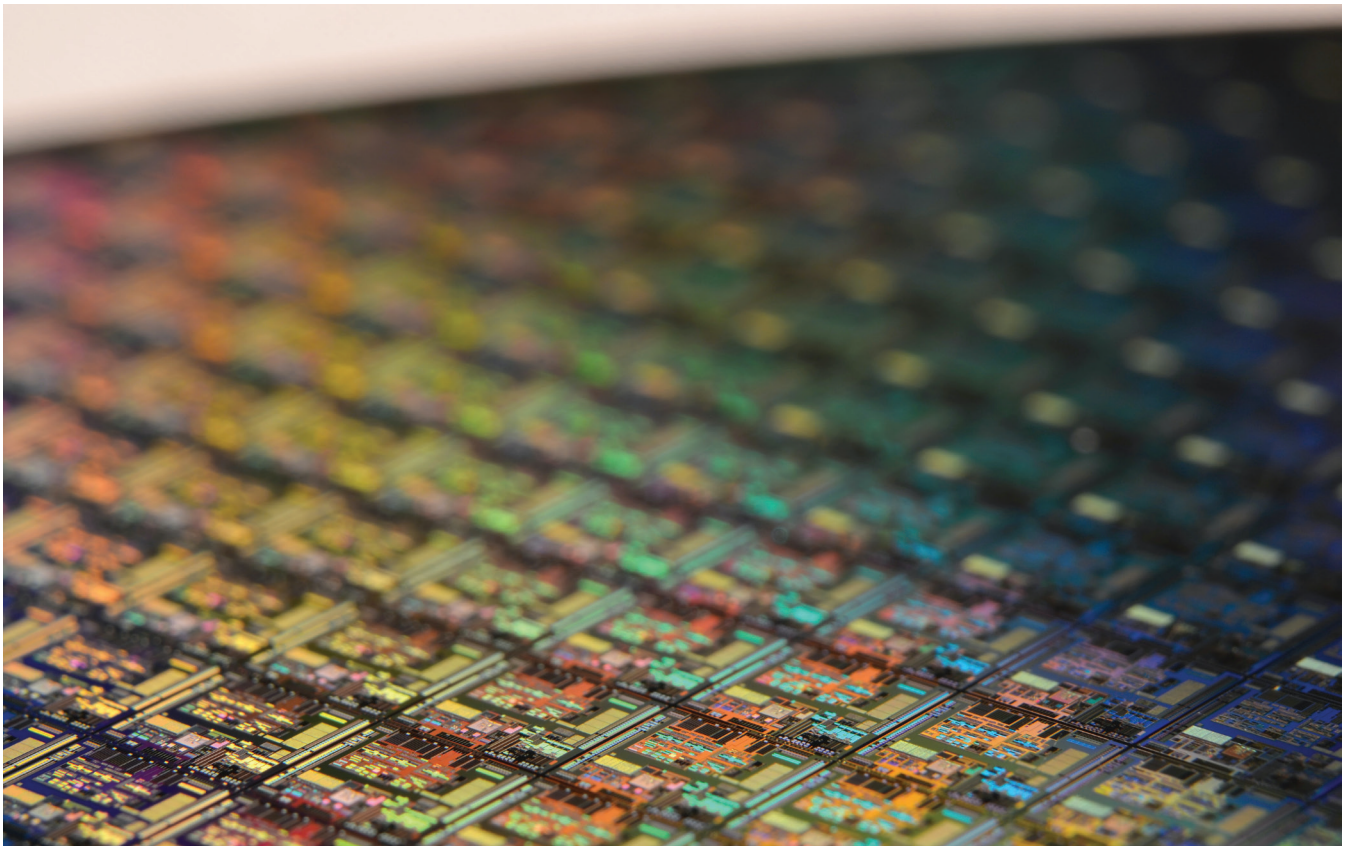
www.EVGroup.com

INTEGRATION SOLUTIONS FOR NEXT-GEN CO-PACKAGED OPTICS

- Enabling customers to manufacture scalable CPO architectures from R&D to HVM
- Hybrid Bonding for low-parasitic PIC/EIC integration and HBM manufacturing
- Heterogeneous integration for laser integration, advanced photonics, and thermal-management materials
- Nanoimprint Lithography for efficient optical coupling and beam shaping
- Advancing next-gen optical I/O and 3D photonic integration



GET IN TOUCH to discuss your manufacturing needs
www.EVGroup.com



Demands for higher productivity, equipment reliability, and sustainability are driving a global shift to predictive maintenance

With rising chip demand and advanced device architecture, fabs face unprecedented complexity in production. The need for higher productivity, increased equipment reliability, and sustainable solutions is driving a global shift toward Smart Manufacturing.

BY DAVID DUVAL, EDWARDS VACUUM, BUSINESS TRANSFORMATION TECHNOLOGY PROGRAMME MANAGER

PREDICTIVE MAINTENANCE, a crucial component of smart manufacturing, integrates extensive sensorization and advanced analytics to allow manufacturers to shift from reactive repair strategies and conservative periodic maintenance to proactive asset management. In an industry where uptime and cost efficiency are critical, these capabilities enable manufacturers to maximize equipment availability, reduce operational risk, and lower maintenance costs.

The sub-fab is a critical enabler of semiconductor manufacturing, housing the vacuum and abatement systems required across most process steps. Failures within the sub-fab can disrupt production and significantly impact fab productivity. While integrating fully connected sub-fab operations through a consolidated global fab control room remains an aspirational objective, the adoption of predictive maintenance strategies is already delivering measurable improvements.

Edwards has developed comprehensive solutions to support the management of sub-fab vacuum assets. Machine learning and artificial intelligence are playing a central and fast-growing role in these initiatives, enabling enhanced operational efficiency, unplanned failure reduction, and ultimately cost reduction. Traditional preventive maintenance relies on fixed service intervals based on historical averages rather than actual equipment condition. To minimise the risk of unexpected

failures, maintenance intervals are usually set shorter than necessary, providing a comfortable safety margin but resulting in the premature servicing of assets that could have continued operating reliably. The gap between the prescribed maintenance interval and the equipment's true remaining useful life represents wasted value, driving up maintenance costs.

Equipment lifetimes can vary widely, and targeting zero failures is not economically feasible: maintenance would become excessively frequent. Predictive maintenance offers a more effective approach in which service is performed only when approaching failure is indicated by relevant signals, enabling equipment to be serviced or replaced near its actual end of life while still avoiding catastrophic failures.

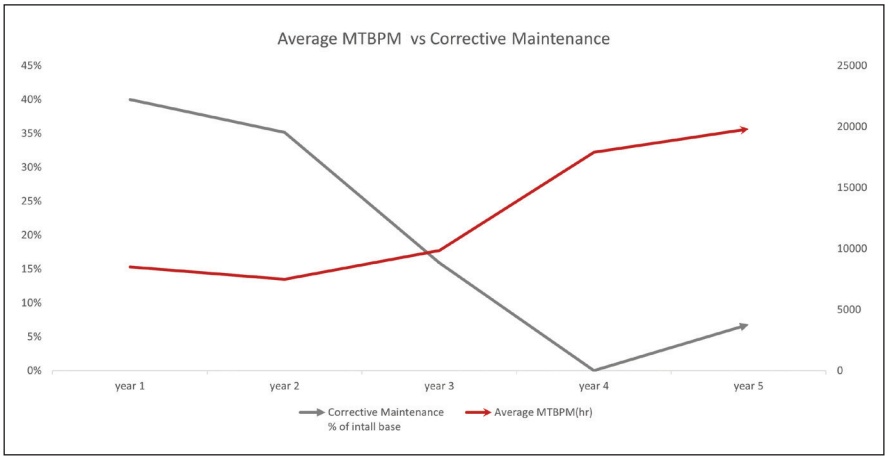
While detecting simple indicators of impending failure is sometimes possible, developing robust models that accurately monitor complex signals to detect degradation and predict failures across a diverse set of assets remains a significant challenge. Rudimentary rule-based monitoring approaches can be effective for some failure modes but generally fail to reliably detect more subtle degradation mechanisms, generating excessive missed failures or false alarms. Increasing sensor coverage, like vibration sensors, may seem an effective way to improve sensitivity and accuracy, but in practice it often results in more signals with poorer signal to noise ratio, raising computational and infrastructure costs. The objective is not to collect the maximum possible amount of data, but to identify the most informative signals.

Effective predictive maintenance therefore depends on addressing three key challenges:

- Collecting the relevant data at the appropriate frequency
- Identifying meaningful signal patterns that indicate an impending failure
- Validating model predictions to confirm performance

Edwards addresses each of these challenges through a targeted approach:

- Collecting only the data required for each specific use case, at the appropriate sampling rate, drawing on decades of vacuum expertise.
- Combining contextual and historical information, domain expertise, and advanced data science techniques to identify complex equipment degradation patterns.
- Physically dismantling and analysing equipment returned to Edwards Service Centers to validate model predictions against actual component condition and failure mechanisms.



➤ Figure 1. Extended MTBPM intervals and corrective maintenance events at ST over time.

Together, these capabilities provide Edwards with a leading predictive maintenance solution that identifies equipment failure mode signatures with a high degree of accuracy.

Production implementation

The effectiveness of predictive maintenance ultimately depends on its ability to deliver measurable improvements in a production environment. The following example illustrates how these principles were applied at scale.

One of our European customers, a leading manufacturer with multiple wafer fabs and assembly/test facilities, approached Edwards to explore a smarter way to manage their extensive pump fleet. They faced a critical issue: process tool downtime caused by unexpected pump failures.

These failures disrupted production, increased costs, and strained resources. Traditional maintenance strategies, whether reactive fixes or scheduled preventive replacements, did not reduce unplanned downtime nor optimize asset utilization.
The initiative began as a collaboration with Edwards regional teams, supported by technical specialists and data scientists in the field and at the factory. Edwards' predictive maintenance platform for vacuum assets was deployed, and vacuum equipment data was collected.

The initial analysis identified consistent patterns preceding failure and models were trained to raise alarms when a complex signature of parameters was identified. Model accuracy was then confirmed through physical inspection of returned equipment at Edwards Service Centers.

With a reliable solution validated for production deployment, predictive models were implemented to prevent unplanned failures and optimize preventive maintenance schedules.

The result was a reduction in vacuum-down events and corrective maintenance interventions, extended mean-time-between-preventive-maintenance (MTBPM) intervals, and a significant improvement in tool availability.

- Unplanned events dropped significantly, from being the majority to just 10% of total replacements in the last two years. In years 1 and 2, the fab averaged 10 Corrective Maintenances (CMs) per year; by years 4 and 5, this was reduced to just 1.5 CMs per year, with one of those caused by external factors.
- 50% reduction in preventative pump replacements. Asset utilization improved from around 10,000 hours MTBPM to around 20,000 hours from year 4 onwards, effectively



reducing preventive replacements by half.
- Both unplanned and preventive replacements decreased simultaneously – a rare achievement in maintenance strategies, proving the effectiveness of predictive maintenance.
- Cost savings achieved by reducing unplanned events, improving predictability, and empowering teams to make proactive, data-driven maintenance decisions.

These results illustrate how predictive maintenance can give manufacturers greater control of subfab operations at scale, improving operational efficiency, asset utilisation, and process tool availability.

Conclusion

Edwards offers monitoring and predictive maintenance plans as separate programs. EdMonitor provides fleet-scale monitoring to capture, display, and store runtime data from vacuum pumps, abatement systems, and other industrial equipment.

EdPredict analyses the data collected by EdMonitor using advanced machine learning and artificial intelligence techniques to predict failures before they occur. As illustrated above, both solutions are already deployed in production across multiple semiconductor manufacturing sites.

Continuous development efforts focus on improving prediction accuracy and expanding coverage across a broader range of equipment.

As part of this roadmap, Edwards is investing in edge computing capabilities that enable predictive models to access more parameters at higher sampling frequencies without overloading the underlying infrastructure.

In parallel, active efforts are underway to explore the benefits of connected equipment, enabling meaningful data sharing across assets to further enhance predictive performance and operational efficiency.

Through continuous innovation, Edwards remains at the forefront of predictive maintenance technology and a key contributor to the advancement of smart manufacturing.

Details of this study are available on the Edwards web site buy scanning the QR code.



Introducing the Swagelok® ALD7C Ultrahigh-Purity Diaphragm Valve.
Built to meet the demands of advanced ALD processes.



Future-Proof ALD Processes.
Precision at Every Pulse.
www.swagelok.com/ALD7C

Swagelok®

All service marks and trademarks shown are owned and registered by Swagelok Company.
© 2026 Swagelok Company. swagelok.com

YOUR POWER PRODUCTS ARE EVOLVING
ARE YOUR TESTERS?

NEXT GENERATION TESTING FOR NEXT GENERATION **3.3 KV SIC PRODUCTS**



COSMIC
SEMICONDUCTOR SOLUTIONS

DISCOVER MORE ON
COSMIC-GROUP.COM/EQUIPMENT



Available is not in control

Balancing output, quality, and risk in high-volume fabs.

BY BORIS SHTEINBERG

A DEPOSITION CHAMBER drops out of production at 2 a.m. Forty minutes later it is back, the dashboard flips from red to green, and two people want opposite things. The operations-center operator has a step starving downstream and wants the tool dispatched now. The module's shift group leader wants test wafers run and confirmed on target before a single product lot goes near it. Both are right. Someone has to choose in the next few minutes, on partial data, with a clock running.

That small standoff is the real work of high-volume manufacturing, and it repeats across a fab every shift. It turns on a distinction the dashboard hides.

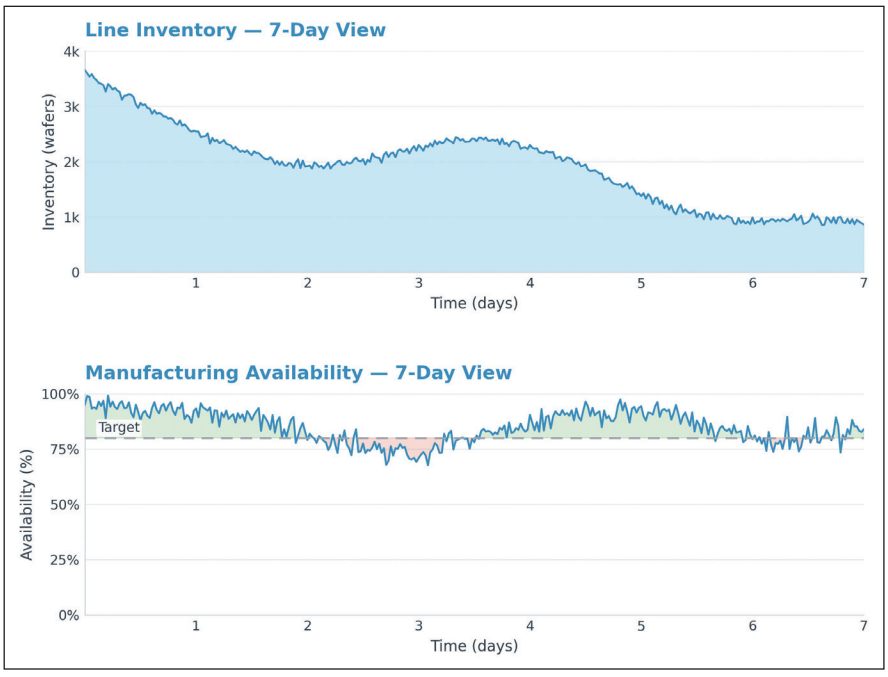
Available is the door being open. In control is whether it is safe to walk through: whether you can trust the tool to do to the next thousand wafers what it did to the last, stable and predictable, within a risk you have actually accepted rather than merely hoped for. A green board is a hypothesis, not a guarantee.

Every fab runs on two clocks. One counts output now: availability, utilization, throughput, the commit for the shift. It is carried by operations-center operators, shift supervisors, shift managers, and the manufacturing engineers keeping product moving.

The other counts consequences later: process stability, equipment health, quality. That clock belongs to the module and process engineering teams, the shift group leaders, the engineering group leaders, who know that a tool released too soon, or a PM deferred to make a number, can send a quiet excursion downstream, one you do not catch until two hundred wafers have already passed through it.

Neither clock is wrong. That is the point. The tension is not dysfunction; it is the system working as designed. Output and control are meant to compete, because a fab that served only one of them would quickly fail at the other.

It helps to be precise about what the green light actually claims. SEMI E10 gives the industry a shared vocabulary for equipment states: productive, standby, engineering, scheduled and unscheduled downtime, non-scheduled time, and from those states the metrics every fab lives by: availability, utilization, mean time between failures, mean time to repair. The vocabulary is exact, and that exactness is what exposes the gap. A tool counted as available may be genuinely productive or merely idle in standby. A tool pushed toward



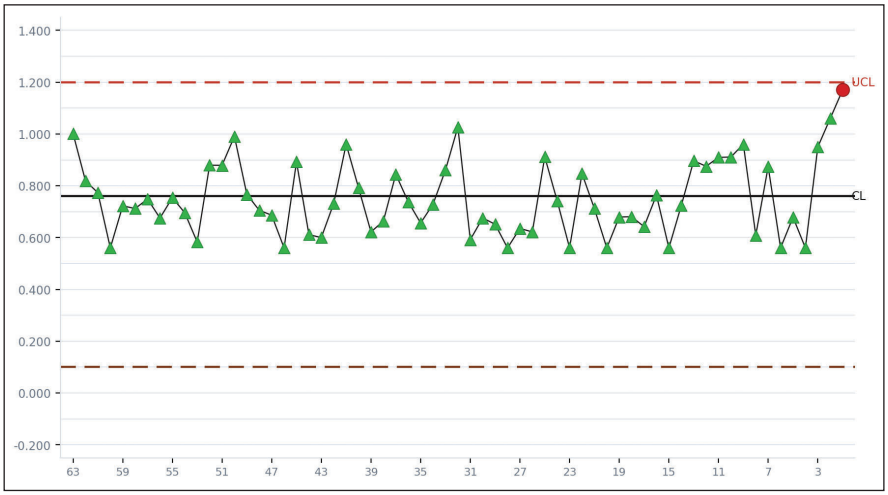
➤ Figure 1. The output-and-availability view: line inventory drawing down over the week, and manufacturing availability swinging around its target.

higher utilization buys throughput today at the cost of cycle time and flexibility tomorrow, because a line run too close to saturation amplifies its own variability.

The states tell you what a tool is doing and how often it does it. They do not tell you whether to trust it with your next lot right now, and they do not tell you whether the number you are optimizing is the one that matters tonight. That is why the same standoff takes so many shapes. A PM comes due, and taking the tool down costs output today while deferring it risks a bigger failure tomorrow. A chamber runs at the edge of spec under a green light, and the supervisor who owns tonight's number is not the engineer who will own the excursion. A tool comes back at partial capability, and someone has to decide whether half a tool now beats a whole tool in three hours. None of these has a clean answer in a procedure. Each is a bet placed with incomplete information, and the standard that names the states cannot place the bet for you. It is worth being concrete about what that looks like on a control chart, because the gap between available and in control becomes almost literal there. Take a parameter the chamber owns — film thickness, etch rate — plotted lot after lot against its limits.

Every individual point can sit inside the three-sigma control limits, so by the single-point rule the tool is in control, productive, green. But a control chart says more than whether the last point passed. Its zone and trend rules exist to catch a process losing control before any one point crosses the limit: readings pushing into the outer zone near three sigma, a Z3 or -Z3 flag, or a run of six or seven points climbing steadily in one direction. That pattern is not a failure yet. It is a forecast of one — a bad trend that says, with the statistics behind it, that an out-of-control point is coming on a tool that is still dispatching product right now. The same logic holds for defectivity: particle or defect counts climbing through their zones, a Z3 on the high side, flagged and trending while still under the limit.

This is exactly where demand decides the call. When the week is tight and the line cannot afford to lose the chamber's availability to make the commit, that trend is easy to read as a problem for later rather than now. The flag fired; someone saw it; and the tool is left



➤ Figure 2. The control view: a parameter holding steady, then its last points trending toward the upper limit while the tool still reads available.

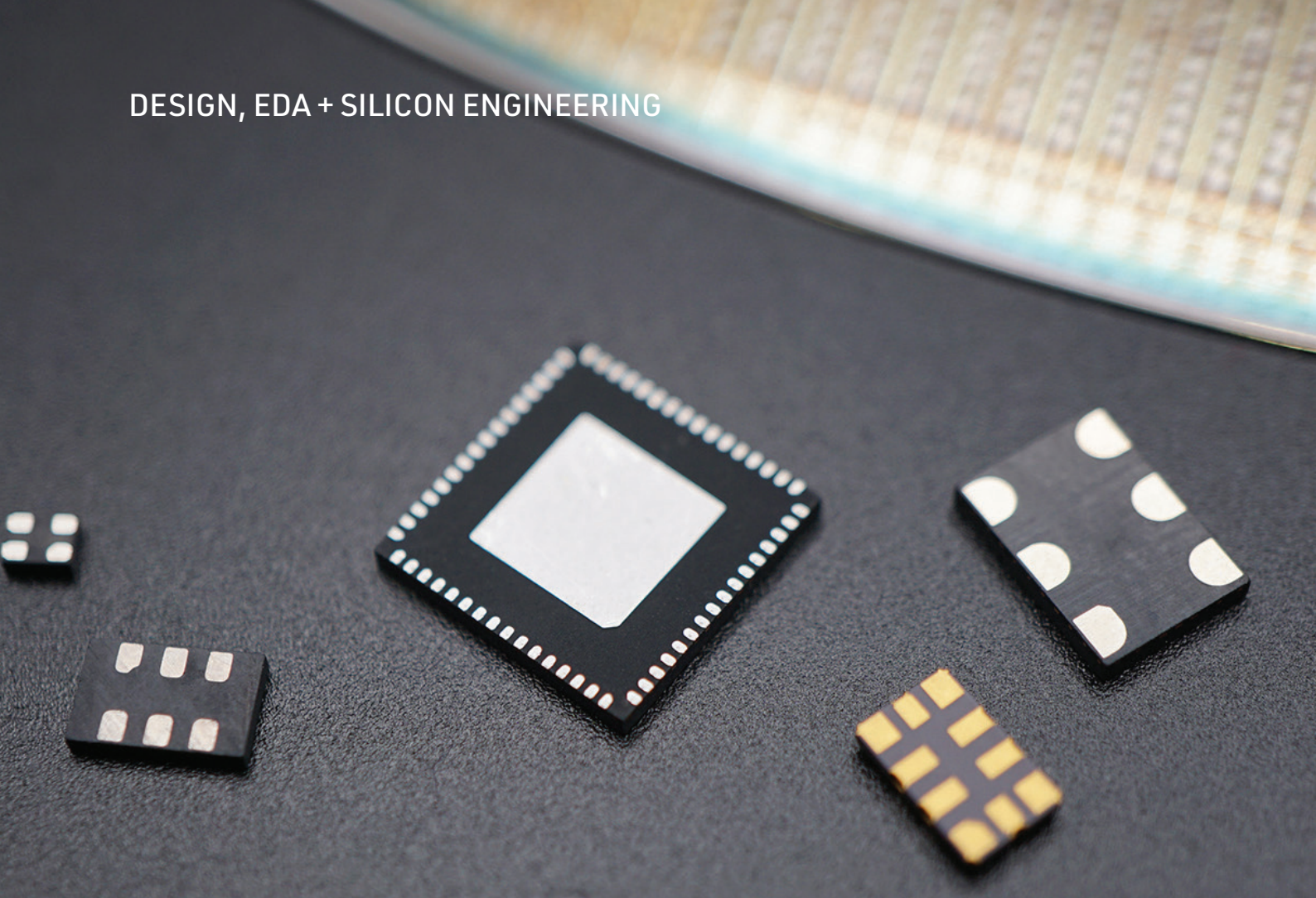
up to production anyway, carried on its availability number while the SPC quietly argues the other way. None of this is a procedure being broken. It is a deliberate bet — that the forecast is wrong, or that the excursion, if it comes, will cost less than the downtime avoided this shift. Sometimes the bet pays and the trend flattens on its own. Sometimes two hundred wafers have already gone through before the point finally lands outside the limit and the board, at last, turns red. The honest version is rarely that the trend was missed. It is that the trend was seen, weighed against the commit, and accepted — which is the whole point.

It helps to be precise about what the green light actually claims. SEMI E10 gives the industry a shared vocabulary for equipment states: productive, standby, engineering, scheduled and unscheduled downtime, non-scheduled time, and from those states the metrics every fab lives by.

The tool was available the entire time. It was never quite in control.

That call lives in people: the engineer who knows this chamber's tells, the operator who has seen this signature before, the group leader who decides which risk the shift will carry. It is built from pattern memory and hard-won feel as much as from data, and much of it is invisible. No one is credited for the excursion they prevented or the downtime they avoided by holding a tool one hour longer. The save never shows up on the wall of screens, which means the people best at this are often the hardest to see, and the easiest to lose when they leave.

This is also the part that does not automate away. We can predict tool health, optimize dispatch, and model utilization better every year, and we should; the floor needs every bit of that signal. But software can recommend; it cannot own the consequence. The decision still belongs to people, and it is rarely one person's to make alone: the operator, the supervisor, the group leader, and the engineer each hold a piece of the picture, and a good call is usually the one they reach together. So the real edge is not a sharper metric or a smarter system. It is people who can weigh availability, quality, equipment health, and risk in a single judgment, made in cooperation rather than in defense of their own number. Output and control were never meant to be enemies; they are two halves of the same call. The fabs that win the next decade will be the ones whose people can hold that balance.



Precision timing: unlocking efficiency and advancing semiconductor sustainability



As AI infrastructure, data centres and edge computing continue their rapid scale-up, the systems that win on efficiency will be those in which timing shapes architecture from the outset of design. Sustainability in the semiconductor industry is often a natural consequence of better engineering.

BY TAYLOR ASTON-NIELSEN, SITIME

THE GLOBAL sustainability drive is in full swing. Against a backdrop of environmental targets, a spotlight has fallen on the semiconductor industry. Its role as the foundational enabler of technology, its exponential growth and its resource intensity means the sector is under pressure to lead and deliver.

As demand from applications such as AI infrastructure, data centres, and edge computing continues to rise, power consumption is becoming a defining challenge. Engineering-led sustainability is taking shape in the semiconductor industry, with meaningful reductions in energy use across entire systems being achieved thanks to the right design choices. At system level, Precision Timing is emerging as a critical factor in improving energy efficiency.

Energy efficiency as a first-order design parameter

The new generation of timing products is being designed in response to system-level efficiency problems. Energy efficiency is now an explicit goal of timing accuracy, rather than being treated as a secondary characteristic or byproduct.

A clear illustration of efficiency as a design-stage priority - rather than an afterthought – is found in AI data centre infrastructure. Every second a GPU sits idle is power spent for zero revenue, with a negative impact on performance per watt (PPW).

With the onus on engineers to design for efficiency, timing has joined compute, network and power delivery to become the fourth critical domain

for optimal GPU utilisation. Better time synchronisation enables improved scheduling and traffic management in AI data centres, which in turn means that the XPUs are spending more time working and less time idling. Work measurements of tokens processed and tokens per watt and dollar are key factors in AI infrastructure buying decisions.

For semiconductor professionals, the practical implication is that more energy efficiency considerations have moved into the early stage of the design process, down to component-level decisions such as timing architecture.

How Precision Timing improves performance per watt

Timing components such as oscillators, resonators and clock ICs are sometimes described as the 'heartbeat' of an electronic system. Every coordinated action across a distributed system depends on a clock signal that all parts of that system can agree on. When that signal drifts, the consequences ripple outwards, causing unreliable data transfer and misaligned subsystems.

A more accurate clock signal shortens that active window. The system spends less time drawing peak current and returns more quickly to a low-power state. Accurate time synchronisation also allows engineers to identify the cause of delays, such as compute imbalance or network congestion. Tighter and more precise synchronisation is eliminating the need for conservative guard bands and wait times at every epoch barrier. Multiplied across millions of cycles and millions of deployed units, efficiency gains compound into a system-wide reduction in energy consumption.

With GPU utilisation in some data centres often operating well below their theoretical capacity, Precision Timing has become a key enabler of higher performance per watt and improved data centre efficiency. Timing components based on Microelectromechanical System (MEMS) technology have achieved new levels of accuracy. The latest TCXO oscillators have taken synchronisation across an AI cluster from 1-10 microseconds to a matter of nanoseconds, or sub-nanosecond in some cases.

The far-reaching impact of timing architecture

Beyond PPW, Precision Timing is playing a part in architecture-level sustainability gains. Timing architecture has moved to the top of the priority list for designers of hyperscale systems such as AI supercomputers, cloud infrastructures and the largest online consumer platforms.

Designs that rely on multiple fixed-frequency oscillators to meet varying frequency, voltage and stability requirements involve an investment in more materials, occupy more board space and increase the manufacturing footprint with each part. MEMS-based timing devices replace several fixed-function components with a single, software-configurable alternative. The sustainability benefits are cumulative; a reduction in the number of devices on the board, an improvement in supply chain efficiency and the simplification of design.

The advantage of programmability in MEMS devices extends into the operational lifetime of products. There are more opportunities for tuning, such as adjusting jitter parameters for a new protocol or compensating for drift.

Smaller footprint paving the way for more efficient designs

In electronics, component miniaturisation is an age-old design objective. Today's dense data centre environments, edge devices and AI infrastructure have become constraints on system efficiency. Every millimetre saved on a timing component can enable an improvement.

MEMS-based timing solutions are resource savers in smaller, more efficient designs, with a footprint that is twice as small as traditional timing devices. MEMS resonators, with up to 3000x less mass, are at the core of this benefit. This reduction in mass translates directly into smaller packages for timing chips.

MEMS timing devices can also bring supporting circuitry on chip. For example, integrating terminal resistors to eliminate external source-bias resistors, offering an internal voltage regulation filter, and negating the need for external load capacitors. The result is fewer external components surrounding the device, shortening the bill of materials and freeing up precious board real estate.

More opportunities for integrations can offer additional power efficiency. Programmable output voltage and dynamic frequency control allow MEMS timing devices to be tuned more precisely to application needs. Features such as switchable output clocks allow unused outputs to be powered down rather than left idling.

Crucially, none of these efficiency gains come at the expense of reliability. Along with miniaturisation, MEMS timing devices also offer superior resilience to shock, vibration and rapid temperature fluctuations, making them more robust in demanding deployment environments such as data centres. Oscillators built on solid-state silicon MEMS technology can be relied on to maintain frequency stability and deliver reliable, accurate timing.

Where desired performance and sustainability meet

Compared to sectors that tend to view sustainability as a list of constraints, the semiconductor sector is in a promising position.

The boom in AI data centres – which will create a semiconductor market worth $1.2 trillion (£895 billion) by 2028, according to the Semiconductor Industry Association – has placed demands on modern timing chips to support improved GPU performance in the face of environmental extremes. In meeting these demands and enabling more performance from each watt consumed, the result is a new level of efficiency and a range of sustainability benefits.

Precision Timing is an enhancement that enables engineers to design lower-power, longer-lasting products while reducing system component counts. Smaller, more resource-efficient devices are translating directly into lower system-wide energy consumption and emissions.

As AI infrastructure, data centres and edge computing continue their rapid scale-up, the systems that win on efficiency will be those in which timing shapes architecture from the outset of design. Sustainability in the semiconductor industry is often a natural consequence of better engineering.

CEA-Leti advances alternative computing architectures

Two recent developments from CEA-Leti highlight the breadth of innovation occurring beyond conventional CMOS scaling. The first demonstrates a hybrid memristor-spintronic Ising machine capable of solving combinatorial optimisation problems using intrinsic annealing mechanisms. The second expands collaboration with quantum start-up Quobly to accelerate the development of scalable silicon spin-qubit technologies through access to Europe's FAMES pilot line infrastructure.

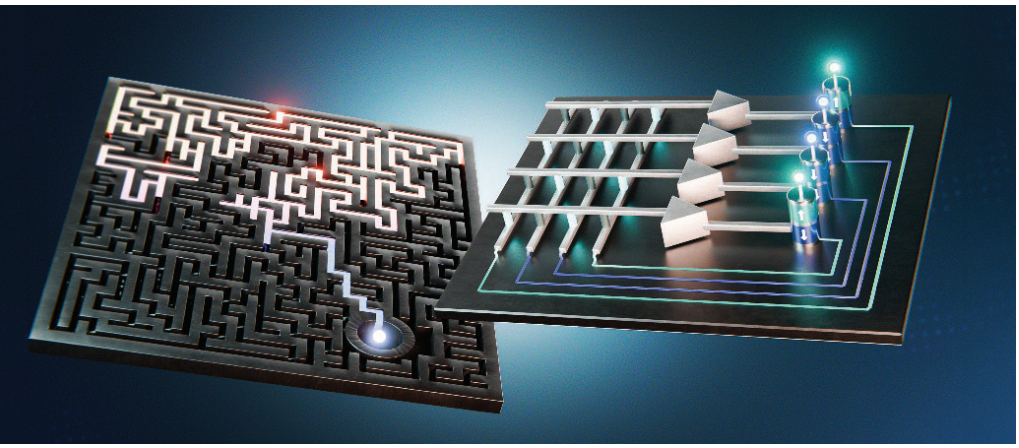
AS PERFORMANCE gains from conventional processor scaling continue to slow, increasing attention is being focused on domain-specific computing architectures capable of delivering significant improvements in performance-per-watt for targeted workloads. CEA-Leti's latest research activities illustrate two distinct but complementary approaches to this challenge: hardware-native optimisation accelerators based on emerging nanoelectronic devices, and silicon-based quantum computing platforms that leverage established semiconductor manufacturing ecosystems.

Recent announcements from the Grenoble research institute showcase advances in both fields. Researchers from CEA-Leti, Spintec and C2N have demonstrated a hybrid nanoelectronic Ising machine that combines hafnium-oxide memristors and stochastic magnetic tunnel junctions (SMTJs) within a unified optimisation architecture. At the same time, CEA-Leti and quantum computing company Quobly have strengthened their long-standing collaboration around silicon spin-qubit technologies, with Quobly gaining access to the FAMES Pilot Line to evaluate future process technologies.

Together, the developments underscore the growing importance of heterogeneous computing approaches as the industry seeks alternatives to traditional von Neumann architectures.

Hardware-native optimisation

Combinatorial optimisation remains one of the most computationally intensive classes of workloads in both scientific and industrial computing. Applications ranging from logistics optimisation and power-grid management to EDA workflows and data-centre resource scheduling routinely involve solution spaces that grow exponentially with problem size.



The hybrid Ising machine reported in Nature Communications addresses these challenges through a hardware implementation of local-search optimisation functions. Rather than executing optimisation algorithms through a sequence of software instructions, the architecture exploits the intrinsic physical behaviour of nanoscale devices to perform key computational operations directly within the hardware fabric.

The system combines hafnium-oxide memristors arranged in crossbar structures with stochastic magnetic tunnel junctions whose thermally driven switching behaviour introduces controlled probabilistic dynamics into the optimisation process. Memristors encode the coupling coefficients that define the optimisation problem, while SMTJs represent the binary state variables whose evolution drives the search towards lower-energy solutions.

A notable contribution of the work is the demonstration of an intrinsic annealing mechanism. Annealing schedules are fundamental to Ising-based optimisation because they govern the transition from exploratory to convergent system behaviour. In many existing implementations, this requires supervisory circuitry and significant control overhead. In the CEA-Leti architecture, however, the coupling between the memristor

network and SMTJs enables annealing behaviour to emerge naturally through adjustment of the memristor read voltage, substantially reducing control complexity.

Toward scalable ising accelerators
Experimental validation was performed on benchmark graph optimisation problems, including weighted MAX-CUT and graph-colouring tasks. The prototype consistently identified optimal solutions for instances comprising 24 vertices in weighted MAX-CUT and 10 vertices in three-colour graph-colouring configurations.

While the present implementation relies on sequential updates and external feedback loops, the underlying architecture is designed with large-scale integration in mind. Future implementations could exploit direct coupling between crossbar arrays and spintronic devices, eliminating analogue-to-digital conversion bottlenecks and enabling highly parallel operation.

From a semiconductor integration perspective, the approach is particularly attractive because both ReRAM-based memristors and magnetic tunnel junctions are compatible with CMOS back-end-of-line processes. This creates potential pathways towards monolithic three-dimensional integration, a key requirement for practical deployment of specialised optimisation accelerators.

For high-performance computing and data-centre applications, such architectures could ultimately provide low-power acceleration of scheduling, placement, routing and resource-allocation workloads that increasingly impact operational efficiency.

Silicon quantum computing moves toward industrialisation

Alongside its work in emerging optimisation architectures, CEA-Leti is continuing to advance Europe's silicon quantum computing ecosystem through its collaboration with Quobly.

Unlike superconducting and trapped-ion quantum computing approaches, silicon spin-qubit technologies seek to leverage mature CMOS manufacturing processes and existing 300 mm semiconductor production infrastructure. This compatibility is



widely regarded as one of the most promising routes towards achieving the device densities and manufacturing scalability required for fault-tolerant quantum processors.

Quobly's QSOI® architecture combines silicon spin qubits with FD-SOI process technology and is being industrialised in partnership with STMicroelectronics. Through access to the FAMES Pilot Line, Quobly will now be able to investigate advanced process modules, lithographic approaches and fabrication flows that may support future generations of quantum processors.

From a semiconductor integration perspective, the approach is particularly attractive because both ReRAM-based memristors and magnetic tunnel junctions are compatible with CMOS back-end-of-line processes. This creates potential pathways towards monolithic three-dimensional integration, a key requirement for practical deployment of specialised optimisation accelerators.

The collaboration builds upon more than 15 years of joint research between CEA-Leti and CNRS, including the demonstration of an industrially fabricated silicon spin qubit on a 300 mm FD-SOI wafer in 2016. The current focus is centred on developing the integration technologies, control architectures and manufacturing methodologies needed to support large-scale quantum processor implementations.

Taken together, the two announcements illustrate CEA-Leti's strategy of pursuing multiple post-CMOS computing pathways simultaneously. While the hybrid Ising machine addresses near-term opportunities for accelerating optimisation workloads through specialised hardware, the Quobly collaboration focuses on the longer-term objective of scalable quantum computing built on established semiconductor manufacturing foundations.

For the semiconductor industry, both initiatives highlight a broader shift towards heterogeneous computing architectures in which emerging device technologies, advanced integration schemes and application-specific accelerators are expected to play an increasingly important role alongside conventional CPUs and GPUs. As AI, HPC and data-centre workloads continue to grow in scale and complexity, such specialised platforms may become critical components of future computing infrastructures.



Before the first weld

What fab gas-line teams should write down when selecting enclosed orbital welding equipment

BY APRIL PEI, MARKETING DIRECTOR, HUBEI FEIYIDE INTELLIGENT EQUIPMENT CO., LTD. (FYID-FEIYIDE)

ULTRA-HIGH-PURITY (UHP) gas lines in a semiconductor fab are built from small-diameter, thin-wall 316L stainless tubing. The joints are almost always made by autogenous orbital GTAW inside an enclosed (closed-chamber) weld head. Most debate focuses on the weld itself: purge quality, oxide colour, internal bead profile. Many problems start earlier, when a team picks weld heads and a power source from a catalogue range instead of from the joints it actually has to make.

This note sets out a short, practical list of what to record before selecting equipment. The worked example is FYID-Feiyide's FXT20 power source with C-series enclosed heads. The same questions apply to any enclosed-head system.

List every OD and wall, not the family range

An enclosed-head family is often described by its total span. The C-series, for example, covers 6.35 mm to 168 mm OD across six heads (C5, C10, C40, C80, C120 and C170). That combined figure hardly matters to a gas-line team. What matters is that each head has its own OD range and wall limit. A tool-hookup package usually mixes 1/4", 3/8" and 1/2" lines with a few larger bulk-gas runs. Put every OD and wall from the isometrics in one table, then map each to a specific head. Gaps and overlaps show up straight away.



➤ FXT20 orbital GTAW power source; stores 200 weld programs and prints a per-weld record.

Measure the clearance, not just the tube

Enclosed heads need radial and axial space around the joint. Next to VCR fittings, valves, regulators, weldments close to a wall or tightly nested tube bundles, the available space often decides which head can be used. It can even decide whether the joint should be made on the bench as a subassembly. Record the clearance at the tightest joints before selecting heads, not after the first fit-up fails.

Decide what weld records QA actually needs

UHP owners increasingly ask for traceable weld records. The FXT20 stores 200 program groups and provides printed welding records as well as USB export of weld data.

The exported fields and file format should be checked against the supplied software version and the owner's record requirements. Confirm the required record format before selecting equipment; USB export alone does not establish compatibility with a site database or quality system.

Keep purge and inspection in the procedure, not the equipment

Equipment ranges do not qualify anything. Purge gas, backing arrangement, oxygen and moisture limits, and internal-surface acceptance come from the owner specification and the WPS/PQR. Operator qualification and the inspection route (visual, borescope, coupon sectioning or whatever the owner requires) stay project-specific. A power source or head range should never appear in a quality plan as evidence that a joint will pass.

Plan a representative mock-up

The simplest risk reducer is a mock-up on the actual tube, fittings and purge set-up, at the tightest clearance on the job. It tests head fit, purge time, program settings and record output together, and it gives the inspector a coupon to judge before production.

A one-page selection record

In practice all of this fits on one page:

- OD and wall for every joint type
- Material and surface finish
- Clearance at the tightest joints
- Purge and backing arrangement from the procedure
- Weld-record format required by QA
- Governing WPS/PQR and inspection method
- Mock-up result

If a team cannot fill in that page, it is not yet ready to choose equipment, whatever the catalogue says.



➤ C-series enclosed orbital weld heads, covering 6.35-168 mm OD across six head sizes.

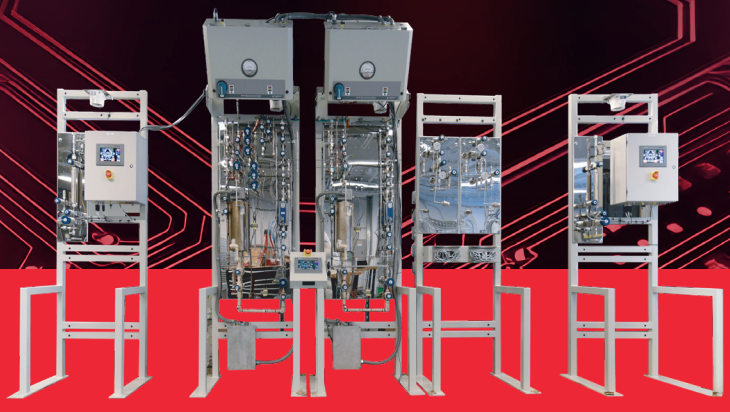
SEMI GAS®

THE FOUNDATION OF SUCCESS

Ultra High Purity Gas Delivery Systems for Semiconductors

Applied Energy Systems' SEMI-GAS® line serves industries with demanding ultra high purity requirements for mission-critical applications.

Our production-ready gas delivery & distribution solutions are globally renowned for quality, precision, and purity. SEMI-GAS® has been the semiconductor industry's source for SEMI S2 compliant gas delivery & equipment for decades, and today's Tier 1 leaders consider SEMI-GAS® their partner of choice.



BROUGHT TO YOU BY

APPLIED ENERGY SYSTEMS

Explore our UHP Gas Delivery Solutions



How California's capital region is emerging as a major chip hub

Barry Broome, President & CEO of the Greater Sacramento Economic Council, outlines the many factors which are contributing to the rise of Sacramento as a globally significant semiconductor hub – not least its proximity to Silicon Valley.

FOR DECADES, Sacramento occupied a distinct place in California's economy. Known primarily as the state capital, an agricultural powerhouse and a strategic logistics centre, it was rarely mentioned in the same breath as Silicon Valley when discussions turned to technology innovation. Yet a quiet transformation has been underway.

Today, Northern California's capital region is rapidly establishing itself as a significant player in the semiconductor industry. Anchored by major investments from companies including Intel, Bosch, Solidigm, Samsung and Micron, and supported by a growing ecosystem of universities, startups, policymakers and economic development organisations, Sacramento is building momentum as one of America's emerging semiconductor centres.

The latest milestone in that evolution came with Bosch's semiconductor expansion in Roseville and the award of $225 million in federal CHIPS Act funding to support the company's operations. But according to Barry Broome, President and CEO of the Greater Sacramento Economic Council (GSEC), the Bosch announcement is merely one chapter in a much larger story.

A combination of market forces, strategic geography, workforce development, collaboration and proximity to Silicon Valley is creating conditions that could make Sacramento one of the most important semiconductor regions in the United States over the coming decade.

From government town to innovation economy

Sacramento's technology aspirations are not entirely new. While the region has traditionally been viewed through the lenses of government, agriculture and logistics, its engineering heritage stretches back decades.

"Sactramento has always been a science community," says Broome. "We've produced excellent engineers for years."

Historically, however, those engineers frequently left.

Graduates from regional universities often saw Silicon Valley as the natural destination for ambitious technology careers. With San Francisco only around 72 miles away, many residents built their professional lives in the Bay Area while maintaining personal ties to Sacramento.
For many years that arrangement appeared perfectly sustainable.

The region enjoyed a relatively affordable cost of living, strong public-sector employment and a comfortable quality of life. There was little perceived urgency to create a globally competitive technology economy.

The 2008 financial crisis changed that perception.



Like many US cities, Sacramento suffered heavily during the housing collapse. Unlike some technology-oriented regions, however, the city lacked a private-sector innovation engine capable of driving rapid recovery.

"The recession lasted longer for us," Broome explains. "Government doesn't necessarily give you the bounce-back effect that other industries can." That experience helped catalyse a strategic shift.

In 2015, the Greater Sacramento Economic Council was formed with a mission to diversify the regional economy, leverage university research capabilities and position Sacramento for growth in emerging technology sectors. Semiconductors would eventually become one of the most important pillars of that strategy.

Rediscovering an overlooked semiconductor heritage

One misconception about Sacramento's rise as a semiconductor cluster is that it emerged from nothing.

In reality, the region has long possessed significant semiconductor foundations.

Intel established a major research and development presence in Folsom during the late 1970s. Hewlett-Packard maintained substantial operations in the broader region. Sacramento also developed expertise through aerospace engineering activity and advanced manufacturing programmes.

Yet much of that activity remained hidden from public view.

Large semiconductor companies often operate with a degree of secrecy that makes their local impact difficult to appreciate.

"Intel was always here, but it was never fully optimised as a regional economic asset," says Broome.

Part of the challenge was visibility. Semiconductor firms focus intensely on engineering, product development and intellectual property protection. They are not always active participants in regional marketing efforts or economic development campaigns.

As a result, Sacramento possessed many of the ingredients necessary for a thriving semiconductor ecosystem without ever fully recognising or promoting them.



Several developments have changed that reality.

The semiconductor supply-chain disruptions experienced during recent years increased awareness of domestic manufacturing and design capabilities. National security concerns elevated semiconductors to strategic importance. Artificial intelligence generated unprecedented demand for advanced computing infrastructure.

Then came the CHIPS and Science Act.

Combined, these factors created an environment in which regions with existing engineering expertise suddenly found themselves in a highly advantageous position.

Sacramento was one of them.

The investments reshaping the region

The growth of Sacramento's semiconductor sector is perhaps best illustrated by the calibre of companies expanding their presence in the region. Among the most significant developments has been Solidigm, the NAND and storage technology company created through SK hynix's acquisition of Intel's NAND business.

Sacramento succeeded in securing Solidigm operations in Rancho Cordova rather than losing the opportunity to San Jose.

According to Broome, the company has become one of the region's most exciting technology success stories.

Its advanced flash storage technologies are increasingly important in AI infrastructure, where massive datasets require rapid, reliable access. That places Solidigm at the heart of one of technology's most important growth markets.

Micron has also expanded substantially, growing its engineering workforce to around 1,000 employees in the region. Samsung has established operations, while a growing number of specialised semiconductor firms have entered the market.

Perhaps most prominent is Bosch.

The German technology giant's investment in Roseville has become one of the defining symbols of Sacramento's semiconductor ambitions.

Bosch is developing a silicon carbide semiconductor operation focused on power electronics applications that support electrification, battery technologies and energy infrastructure. Silicon carbide devices are increasingly important as industries transition toward electric vehicles, renewable energy systems and more efficient power management technologies.

For California, where electrification and decarbonisation remain major priorities, the strategic significance is considerable.

The federal government's decision to support the project through substantial CHIPS Act funding further reinforced confidence in the region's long-term prospects.

Yet Broome believes another announcement may prove equally important.

Intel's recent decision to recommit to its Folsom site, retain real estate assets and pursue additional investment has, in his view, strengthened Sacramento's position as one of America's leading semiconductor design centres.

"I think it solidifies us as a top-five design centre in the United States," he says.

Design, not just manufacturing

While semiconductor manufacturing often generates the headlines, Sacramento's greatest strength may ultimately lie elsewhere.

Broome sees the region emerging primarily as a centre for semiconductor design, research and development.

That distinction matters.

Advanced semiconductor design requires highly skilled engineering talent and creates high-value jobs. It also generates innovation spillovers that support startups, university research and broader economic growth.

The region already hosts nine semiconductor companies pursuing expansion activities, while additional firms are evaluating potential investments.

Recent discussions with Taiwanese semiconductor organisations suggest growing international interest.

According to Broome, several specialised chip-design companies are actively considering Sacramento as a location for research and development facilities.

This aligns with broader trends reshaping the global semiconductor industry.

As chip design becomes increasingly complex, companies are searching for locations that combine access to talent, lower costs and proximity to major technology ecosystems.

Sacramento offers all three.

The Silicon Valley advantage

No discussion of Sacramento's growth would be complete without considering its relationship with Silicon Valley.

Rather than viewing the Bay Area as competition, Sacramento increasingly sees itself as a complementary extension of California's technology engine.

The numbers are compelling. Housing costs remain substantially lower than those found in San Jose or San Francisco. At the same time, travel times between Sacramento and Silicon Valley allow relatively easy movement of people, expertise and ideas.

Broome describes the relationship as a form of nearshoring.

"We're 72 miles away at roughly half the price," he says.

This value proposition has become increasingly attractive as Silicon Valley grapples with rising housing costs and operational expenses.

A substantial workforce exchange already exists between the two regions. Many Sacramento residents work in Bay Area technology companies, while thousands of Bay Area professionals live in Sacramento. Those connections are becoming stronger rather than weaker.

Broome believes Sacramento, San Francisco and San Jose are steadily evolving into an integrated economic corridor where each region performs distinct but complementary functions.

Under that vision:

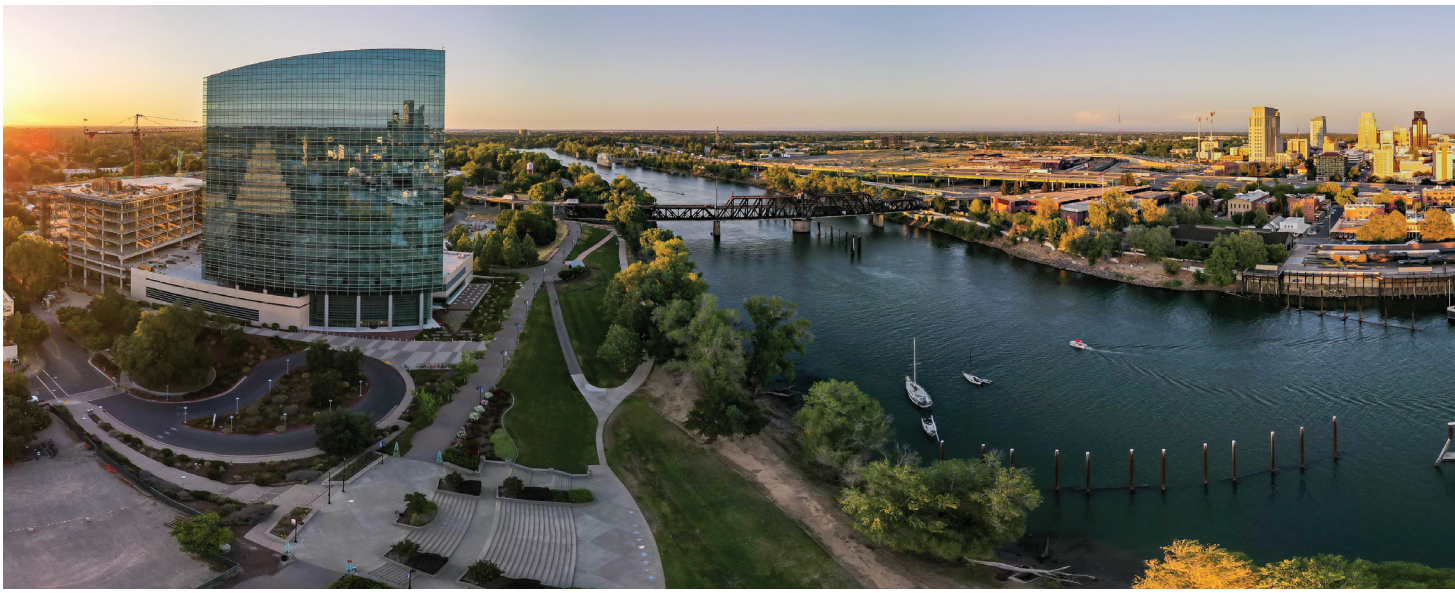
- San Francisco becomes the centre of AI innovation.
- San Jose remains the dominant hardware and semiconductor headquarters location.
- Sacramento specialises in semiconductor design and engineering.

Together, the three regions create a technology ecosystem that few global competitors can match.

Building the next generation workforce

Talent remains one of the most important differentiators in semiconductor development.

Recognising this reality, Sacramento's universities and educational institutions



are increasingly aligning themselves with the industry's needs. Broome identifies workforce development as one of the region's greatest strengths.

A distinctive feature of Sacramento's economic development model is the unusually close collaboration between academia, industry and government. University presidents, community college leaders and business executives all participate directly in strategic planning initiatives.

That structure enables educational institutions to respond more quickly to changing market demands.

Sacramento State's planned engineering school provides a particularly notable example. Rather than simply expanding an existing engineering curriculum, regional stakeholders are asking a more ambitious question: what should a modern engineering school look like in the age of AI and semiconductors? The result could be a fundamentally different educational model designed specifically around future workforce requirements.

The California State University system has already committed significant funding toward the initiative, while private-sector partners are expected to contribute substantial additional investment.

Meanwhile, institutions such as Sierra College continue earning recognition for advanced workforce programmes, including highly regarded mechatronics training.

The objective is clear: create a pipeline capable of supplying semiconductor companies with engineers, technicians and researchers for years to come.

A growing startup ecosystem

Large corporations are not the only source of momentum.

Sacramento's startup environment is becoming increasingly active. Several venture-backed semiconductor and AI companies have emerged from the region, benefiting from the deep engineering expertise developed through decades of Intel and other technology-sector activity.

Broome points to companies such as Blaize and Azimuth AI as examples of growing entrepreneurial strength.

Many of these businesses are being founded by engineers who gained experience inside larger technology organisations before launching their own ventures.

This pattern mirrors the evolutionary pathway followed by countless successful technology ecosystems worldwide.

Experienced talent accumulates within established firms, eventually generating spinouts, startups and entirely new innovation networks.

Foreign direct investment is also playing a growing role.

Strong relationships with South Korean technology firms including Samsung, SK hynix and Solidigm have created opportunities for Sacramento to deepen international technology connections.

Broome believes the city could become an important US destination for Korean technology investment in coming years.

Why collaboration matters

One recurring theme throughout Sacramento's development story is collaboration.

While many regions seek to build semiconductor ecosystems, not all achieve the same degree of alignment between stakeholders.

Broome believes Sacramento's collaborative culture provides a significant competitive advantage.

The region's economic development board includes leaders from higher education, utilities and private industry.

This structure is particularly important given the semiconductor industry's enormous infrastructure requirements.

Energy availability has become one of the industry's most critical concerns, especially as AI data centres drive unprecedented demand for electricity.

By involving utility companies directly in strategic planning, Sacramento hopes to address infrastructure challenges proactively rather than reactively.







This collaborative approach also streamlines decision-making for companies evaluating investments. Rather than navigating multiple disconnected organisations, prospective investors gain access to a coordinated network capable of addressing talent, infrastructure, permitting and workforce questions simultaneously.

As semiconductor projects become increasingly complex, that efficiency can prove decisive.

The Bosch blueprint

The Bosch project also highlights another advantage Sacramento possesses: existing semiconductor infrastructure.

Operational semiconductor facilities are relatively rare assets.

Building a new fab from scratch requires enormous investment, lengthy permitting processes and extensive site preparation.

Bosch benefited from acquiring and upgrading an existing facility in Roseville rather than starting entirely from zero.

That decision potentially saved significant time and capital.

Additional advantages also came into play. The facility was located within the jurisdiction of Roseville Electric, a public power utility able to offer greater flexibility than some investor-owned utility structures.

Roseville itself has developed a strong reputation for business-friendly economic development practices. Combined, these factors created a highly competitive investment proposition.

As Broome notes, economic development involves both strategy and timing.

Sometimes success comes from recognising and maximising opportunities when favourable conditions align.

The Bosch project exemplifies that principle.

Looking toward the future

So what comes next?
Broome is optimistic that Sacramento's semiconductor story is still in its early stages.

The region expects additional technology investments in the near future, potentially including projects valued between $500 million and $1 billion. Longer-term ambitions are even more significant.

Broome believes California could eventually secure another major semiconductor fabrication project worth between $10 billion and $25 billion.

If that happens, he argues Sacramento may be among the strongest candidates in the state.

The region's combination of affordability, workforce quality, infrastructure, university engagement and proximity to Silicon Valley creates a compelling proposition.

At the same time, California must continue addressing competitiveness challenges.

Balancing economic growth with environmental goals, infrastructure planning and regulatory efficiency will remain an important challenge for policymakers. Yet Broome remains convinced that Sacramento can play a central role in California's technology future.

Conclusion

Sacramento's emergence as a semiconductor hub is not the result of a single investment, policy initiative or company expansion.
Rather, it reflects a broader transformation decades in the making. The region is leveraging engineering talent, educational institutions, strategic geography, industrial heritage and growing ties to Silicon Valley to create a new identity within the global technology landscape.

Bosch's Roseville expansion, Intel's recommitment to Folsom, the growth of Solidigm and Micron, and increasing international interest all suggest that momentum is building.

Whether Sacramento ultimately becomes one of the world's most recognisable semiconductor centres remains to be seen.

What is increasingly clear, however, is that California's capital is no longer simply a government town or agricultural centre. It is becoming a serious semiconductor destination in its own right, and a critical component of a broader Northern California technology ecosystem that may prove central to the future of AI, advanced computing and semiconductor innovation.

Photos supplied by Greater Sacramento Economic Council (GSEC).

SCREEN

Driving Innovation Aimed at Net Zero with Our Partners

Go Green toward Net Zero

Provision of eco-friendly products and services

The SCREEN Group is working to expand sales of products that are certified as Green Products according to its independent evaluation standards for excellent environmental performance.



Single Wafer Cleaner
SU-3200



Spin Scrubber
SS-3200



Coat/Develop Track
DT-3000

SCREEN Semiconductor Solutions Co., Ltd.

SEMICON® EUROPA Nov 10–13, 2026, SCREEN Booth #B1700



Accelerating the next era of custom AI silicon

Mohit Gupta, Co-Founder and CEO and Sunil Bhardwaj, Co-Founder of TYLsemi, explains that TYLsemi is the first chiplet platform company to deliver a full portfolio across IO, power delivery, and memory paired with a chiplet-based custom silicon design, integration and supply-chain ownership.

ARTIFICIAL INTELLIGENCE is reshaping the technology landscape at an unprecedented pace. As hyperscalers, cloud providers, sovereign technology initiatives, and AI-native companies race to develop ever more powerful computing infrastructure, the semiconductor industry is facing a fundamental challenge: traditional approaches to chip design are struggling to keep pace with the scale, complexity, and economics of modern AI workloads.

Into this environment steps TYLsemi, a company that has emerged from stealth with significant funding and an ambitious mission. Founded by semiconductor industry veterans, TYLsemi is positioning itself as the industry's first full-stack chiplet platform company, providing customers with a standards-based portfolio encompassing connectivity, power delivery, memory, custom silicon integration, and supply-chain ownership.

According to Founder and CEO Mohit Gupta, the company's goal is simple but transformative: enable organisations to build custom AI silicon faster, cheaper, and with significantly less risk.

The growing demand for custom AI chips

The rise of AI has created a dramatic shift in the semiconductor landscape. Not long ago, only a handful of companies possessed the resources and expertise necessary to develop custom silicon. Today, that picture looks very different.

"Every hyperscaler, every AI company, and even sovereign initiatives are designing their own custom chips," Gupta explains.

The drivers are varied. Some organisations seek performance advantages by tailoring silicon to specific workloads. Others want greater control over their technology stacks or independence from external suppliers. Whatever the motivation, the trend is clear: custom silicon has become a strategic asset.

Yet building advanced AI processors has become extraordinarily difficult. Development costs continue to climb as process technologies move toward 2nm and beyond. Engineering teams face immense challenges related to performance, power consumption, memory bandwidth, connectivity, manufacturing yield, packaging complexity, and time-to-market pressures.

TYLsemi's founding thesis is that companies do not need to reinvent every component of an AI accelerator from scratch.

Instead, they can focus on what differentiates them while leveraging reusable chiplet building blocks for common functions such as connectivity, memory, and power delivery.

That belief formed the basis for the company's creation.

"We identified a gap in the market," Gupta says. "Customers repeatedly told us there was a need for a company that could provide those foundational components while helping them

integrate everything into custom silicon."

Why chiplets have become the new standard

Although chiplets have been discussed for many years, Gupta believes the industry has now reached a critical inflection point.

Chiplets, essentially modular blocks of silicon that can be combined into a larger processor package, offer an alternative to traditional monolithic chip designs. Rather than placing every function on a single, massive die, designers can create specialised chiplets optimised for different tasks and manufacturing processes.

Several powerful forces have accelerated this transition.
The first is physics.

For decades, Moore's Law enabled predictable increases in transistor density and performance. While advances continue, those gains are becoming increasingly difficult and expensive to achieve. At the same time, AI workloads have generated unprecedented demands for compute, memory bandwidth, and connectivity.

"The pain of staying monolithic now exceeds the pain of switching to a chiplet-based design," Gupta argues.
Economics reinforce the trend.
As process technologies move to 3nm, 2nm, and eventually angstrom-scale nodes, non-recurring engineering (NRE) costs have soared. Large monolithic chips are more expensive to design, manufacture, and validate. Yield challenges become increasingly severe as die sizes increase, making defects more costly.

Chiplet architecture offers a way around those limitations.

By breaking designs into smaller functional blocks, companies can improve yields, optimise process node selection, and reuse proven silicon across multiple products.

At the same time, important industry enablers have emerged.

Advanced packaging technologies such as 2.5D and 3D integration have matured significantly and are now deployed in production at leading foundries and OSATs. Just as importantly, the industry has begun converging around open standards.

The critical role of UCIe

One of the most significant developments for the chiplet ecosystem has been the emergence of the Universal Chiplet Interconnect Express (UCIe) standard.

In the early days of chiplets, fragmentation threatened widespread adoption. Major semiconductor companies developed proprietary die-to-die communication methods, creating interoperability challenges and limiting ecosystem growth.

UCIe is changing that dynamic.

"The industry needed something equivalent to what PCI Express and Ethernet achieved for board-level communication," Gupta says.

As chiplets became more common, a standard mechanism for die-to-die connectivity became essential. UCIe provides that common language.

The significance extends beyond technical interoperability. Open standards reduce vendor lock-in, encourage ecosystem growth, and allow customers to adopt multi-source procurement strategies. This capability is particularly valuable in the AI era, where supply-chain resilience has become a major strategic concern.

Gupta sees widespread support from hyperscalers as a powerful validation of the standard's future.

"If organisations with effectively unlimited resources choose to adopt open standards, that tells you something important."

Looking ahead, he believes chiplet ecosystems could eventually resemble the modular flexibility that engineers once enjoyed with traditional PCB-based component design.

The industry is not yet at that point, but the direction is becoming increasingly clear.

Beyond chiplets: the full-stack platform approach

Many companies in the semiconductor ecosystem provide individual IP blocks or specific chiplet technologies. TYLsemi believes that approach is insufficient.

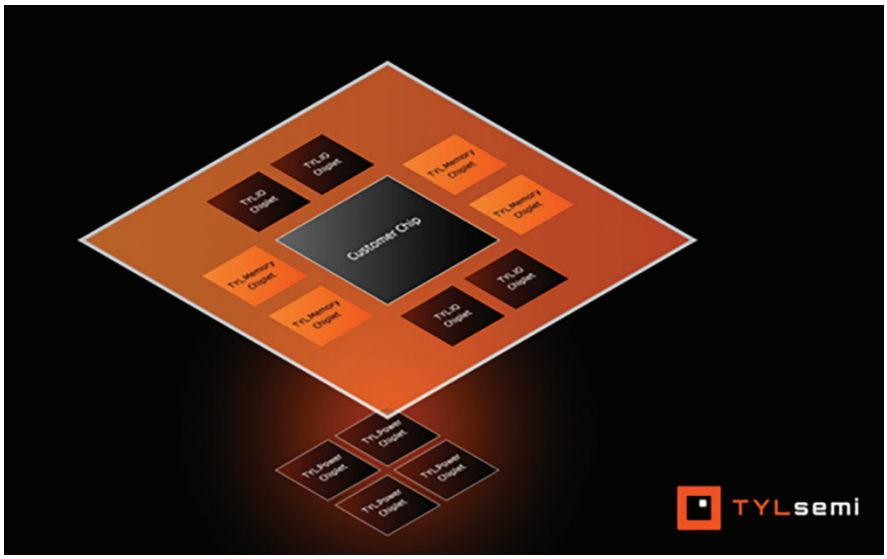
The company's strategy centres on providing an integrated platform spanning three critical dimensions:

- Connectivity
- Power delivery
- Memory

These elements are combined with custom compute integration to create complete silicon solutions.

"When you build silicon, you need to take a system-level view," Gupta explains.

The challenge is that all three technology vectors are deeply interconnected. Decisions surrounding memory architecture influence power characteristics. Connectivity choices impact performance and efficiency.





➤ Sunil Bhardwaj (left) and Mohit Gupta (right).

Power delivery affects both thermal management and compute density.

Optimising one element in isolation often creates constraints elsewhere.

By addressing all three simultaneously, TYLsemi believes it can unlock efficiencies that would otherwise remain inaccessible.

The company's platform model also supports manufacturing flexibility.

Customers increasingly want the ability to work with multiple foundries, packaging partners, and supply-chain providers. By supporting a multi-source approach, TYLsemi aims to reduce dependency on any single manufacturing pathway.

For Gupta, simply offering one category of chiplet would represent only part of the solution.

"A complete platform provides significantly more value than a standalone component."

Reducing development time and cost

Among TYLsemi's most eye-catching claims is the ability to reduce custom AI silicon development costs and timelines by as much as 50%.

Gupta argues that these savings emerge from several sources. Consider a modern AI inference accelerator.

The compute engine may be manufactured on a cutting-edge 2nm process node, while 3nm is close to leading edge. IO chiplets typically sit on 5/6/7nm or older. Power delivery components may utilise 12nm, 16nm, or even older manufacturing processes.

Traditionally, each of these blocks would require separate development programmes, verification efforts, mask sets, engineering teams, and extensive qualification work.

Reusable chiplets remove much of that burden.

The benefits go far beyond the cost of mask sets alone.

Engineering resources can be redirected toward differentiated compute architectures rather than foundational infrastructure components.

Verification cycles are reduced because existing silicon has already been validated. Packaging strategies can be reused. Development risk is lowered. The cumulative impact can be substantial.

Gupta points to recent examples, such as OpenAI's collaboration with Broadcom, where silicon development timelines were dramatically compressed.

"The industry has gone from accepting two-to-three-year chip development cycles to trying to deliver major products annually," he notes.

Under those conditions, reuse becomes a strategic necessity rather than a convenience. The result is a shift from massive engineering programmes involving hundreds of specialists toward smaller teams delivering products in significantly shorter timeframes.

Solving AI's power problem

While AI discussions typically focus on compute performance, Gupta argues that power has become one of the industry's defining challenges. In many cases, data centre expansion is no longer constrained by chip availability.

Instead, the limiting factor is energy. Power consumption figures for next-generation AI accelerators continue to rise sharply. Whereas high-performance CPUs and networking equipment once consumed several hundred watts, future AI processors are expected to approach or exceed multiple kilowatts per device.

This trend creates challenges throughout the infrastructure stack, from grid connectivity and facility design through to board-level and package-level power delivery.

TYLsemi's TYL.Power technology focuses specifically on what Gupta calls the "last mile" of power delivery. Rather than concentrating solely on bringing power to the package, the company's objective is to optimise how energy reaches the transistor level inside the device itself.

This requires greater intelligence and control.

TYLsemi aims to provide granular power management capabilities that allow operators and system designers to optimise energy consumption according to workload requirements. The potential savings are significant.

Gupta estimates that reducing consumption by approximately 300 watts on a 1.5kW accelerator could collectively save ~$10B of dollars across the AI infrastructure ecosystem by the end of the decade. While exact figures will ultimately depend on deployment volumes and energy pricing, the underlying argument is compelling: even small efficiency gains become enormously valuable when multiplied across millions of AI accelerators.

Addressing the connectivity bottleneck

If power is one side of the equation, connectivity is the other.

A growing challenge for semiconductor designers is the widening gap between advances in compute performance and improvements in I/O technology.

Processors are becoming increasingly capable of handling massive volumes of computation, but moving data in and out of those devices remains difficult.

For AI systems, this creates a significant bottleneck. Raw compute capability provides limited value if data cannot reach the processor quickly enough. One challenge is that advanced I/O



technologies often lag behind the latest logic processes. As organisations adopt cutting-edge nodes for AI compute engines, corresponding SerDes and connectivity technologies may not yet be available or fully optimised.

This is where TYLsemi's TYL.IO. family enters the picture.

Designed to support PCIe, Ethernet, and emerging connectivity technologies, the platform is intended to give customers access to high-performance interfaces without requiring them to develop those capabilities independently.

The objective is straightforward: allow customers to focus on their AI differentiation while ensuring that connectivity infrastructure keeps pace with increasingly powerful processing engines.

As AI clusters continue expanding and architectures become more distributed, that capability will play an increasingly important role.

Early traction and industry validation

Although TYLsemi remains a young company, it has already secured engagements with tier-one customers and is working alongside TSMC on production samples. According to Gupta, customer response has been overwhelmingly positive.
In some cases, prospective customers have expressed a desire to begin integrating TYLsemi's solutions immediately.

Of course, semiconductor development timelines remain measured in years rather than months.

The company's first chiplet samples are expected during the second half of next year, while larger customer programmes incorporating those chiplets are targeting deployments around 2028.

Nevertheless, customer engagement extends far beyond product demonstrations.
TYLsemi actively works with customer architects to gather feedback and ensure future products align with long-term roadmaps.

> The success of chiplet-based design ultimately depends on ecosystem participation. No single company can solve all the challenges associated with advanced AI infrastructure.

This collaborative development model is particularly important in a market where semiconductor projects often begin several years before final deployment.

The company's leadership team also brings considerable industry experience.

Although TYLsemi itself is still in its infancy, Gupta emphasises that the organisation has been built by engineers who have previously delivered high-volume, complex silicon products. That experience provides credibility with customers evaluating long-term strategic partnerships.

Building on a strong ecosystem

A key pillar of TYLsemi's strategy is its close relationship with ecosystem partners.

At the centre of that ecosystem sits TSMC, currently serving as the company's foundry partner.
Gupta describes the collaboration as particularly valuable given TSMC's track record of enabling advanced semiconductor innovation while maintaining an open foundry model.

Beyond manufacturing, the company relies on cooperation across the broader semiconductor value chain, including packaging providers, OSAT partners, EDA vendors, and system integrators.

The success of chiplet-based design ultimately depends on ecosystem participation.
No single company can solve all the challenges associated with advanced AI infrastructure.

Instead, progress relies on standards, interoperability, and collaboration across multiple technology domains.
A future built on heterogeneous silicon
Looking ahead, Gupta envisions a future where building custom silicon becomes dramatically more accessible.

His goal is not to help customers establish internal semiconductor companies. Rather, it is to eliminate the need for them to do so.

In that future, organisations focus on the elements that make them unique, whether that is compute architecture, software frameworks, compilers, or AI models. Specialist partners such as TYLsemi would handle the complexities of transforming those ideas into physical silicon. Ultimately, Gupta sees chiplets enabling a truly heterogeneous semiconductor ecosystem.
In this model, individual dies could be sourced from multiple regions, built on different process technologies, and integrated into unified products through advanced packaging techniques.
While significant challenges remain, the industry has already demonstrated an ability to overcome obstacles that once seemed insurmountable.

Technologies such as 3D packaging and advanced die stacking have moved from conceptual research to commercial reality in just a few years. For Gupta, this progress suggests that even more ambitious visions may be achievable.

Conclusion

The semiconductor industry is entering a period of profound transformation. AI is driving unprecedented demand for custom silicon, while simultaneously exposing the limitations of traditional monolithic design methodologies.
TYLsemi believes the answer lies in a modular, standards-based future built around chiplets.

By combining connectivity, power delivery, memory, custom integration capabilities, and ecosystem partnerships into a single platform, the company is seeking to simplify one of the most complex engineering challenges in modern technology.
Whether that vision ultimately reshapes the industry remains to be seen. However, the forces driving chiplet adoption appear increasingly powerful: rising costs, growing design complexity, mounting power constraints, and relentless pressure to accelerate innovation.

As AI infrastructure continues to scale, the ability to build faster, more efficient, and more flexible silicon could become one of the defining competitive advantages of the next decade.

And if Gupta's vision proves correct, future companies looking to build AI chips may spend far less time worrying about semiconductor development itself and far more time focusing on the workloads that make their solutions unique.

Seeing **beyond**

ZEISS

Struggling to understand your real manufacturing costs?

A new KPI connects cost, yield, and financial logic into one consistent metric.

Download Whitepaper

Unlock the value of Manufacturing Cost per Good Die with a scalable, data-driven approach to semiconductor cost management. By ZEISS Digital Innovation.



ANGELTECH

CS INTERNATIONAL CONFERENCE | PIC INTERNATIONAL CONFERENCE | PE INTERNATIONAL CONFERENCE | AP INTERNATIONAL CONFERENCE | AIS ANGELTECH INNOVATE SUMMIT

SAVE THE DATE

12-14 APRIL 2027

Sheraton Brussels Airport Hotel

To find out more about our sponsor and speaker opportunities, contact us today on:
+44 (0) 2476 718970

or email:

info@csinternational.net

info@picinternational.net

info@peinternational.net

info@angel-tech.net



The hidden impact of lubricant selection on contamination control and yield in semiconductor fabs

A technical look at how lubricant chemistry, outgassing behavior, and particulate control influence semiconductor tool performance.

BY BRIAN HOLTKAMP, HEAD OF GLOBAL BUSINESS SEGMENT, SEMICONDUCTOR, FUCHS LUBRICANTS COMPANY

SEMICONDUCTOR manufacturing environments are engineered around exceptional contamination control. Airborne particles, process residues, wafer handling, and tool surfaces all receive close scrutiny. Lubricants, however, can be easier to overlook because they sit inside mechanical assemblies rather than directly in the process itself, but that does not make them any less critical to the success of a fab's yield. A lubricant used in a bearing, linear guide, valve, actuator, robot, or sealing interface can support precision motion and component life while also creating molecular or particulate contamination pathways if its chemistry is not matched to the operating environment.

Those contamination pathways generally fall into two categories. Molecular contamination occurs when volatile lubricant constituents evaporate or outgas, migrate through the tool environment, and condense onto a sensitive surface. Particulate contamination occurs when solid material smaller than a grain of sand is introduced through lubricant manufacturing, handling, thickener agglomeration, degradation, or mechanical wear during operation. Either pathway can affect wafers, optics, sensors, metrology surfaces, and motion systems before the lubricant itself is identified as the source.

Fabs often use shielding, barriers, purge strategies, and guarded assemblies to limit contaminant migration. Those controls reduce exposure, but they cannot compensate for a lubricant whose volatility profile, cleanliness level, or film stability is wrong for the application. Once condensable material or particulate matter reaches the process environment, migration becomes difficult to predict and root cause analysis becomes harder. A trace deposit on an optical surface or a small amount of particle release near a wafer handling mechanism can create yield loss, downtime, and investigation costs that far exceed the cost of stronger selection at the outset.

Vacuum, motion, and heat

change the failure mechanisms

A lubricant that performs well in a conventional industrial environment may exhibit altered lubrication film stability under vacuum exposure, cyclic loading, or thermal cycling. Vacuum changes the threshold for volatilization. Under atmospheric pressure, evaporated molecules are partially constrained by surrounding gas molecules and local pressure equilibrium.

Under vacuum, reduced ambient pressure lowers that constraint and allows low-molecular-weight fractions to leave the lubricant surface more readily. Lubricants with lower vapor pressure, higher molecular weight base oils, and tighter molecular distributions generally reduce the population of volatile species available for release.

Outgassing should therefore be treated as a chemistry and process compatibility issue rather than a generalized cleanliness claim. The concern is not only how much material leaves the lubricant, but whether that material can redeposit as condensable molecular contamination on nearby surfaces. In deposition, etch, metrology, or wafer-handling systems, even low levels of condensable material may affect process uniformity, optical performance, sensor response, or surface cleanliness.

Chemistries such as perfluoropolyethers, multiply alkylated cyclopentanes, and polyphenyl ethers may be considered for demanding vacuum or high-temperature environments because they can provide low volatility and thermal stability. No chemistry is universally suitable, however. Engineers still have to weigh wear protection, elastomer and plastic compatibility, regulatory exposure, service interval, and particle control against molecular contamination risk.

Motion introduces a separate failure mechanism. Bearings, robotic arms, linear guides, valves, and actuators subject the lubricant film to shear, load variation, and repeated cycling. These forces can influence oil separation, migration, film thickness, and particle generation over time. In rolling element bearings, particles larger than the elastohydrodynamic film can interrupt separation between balls and raceways, causing debris denting, progressive wear, and additional particle formation. In precision motion systems, that can reduce repeatability, increase torque variation, shorten component life, and affect nearby wafers or optical surfaces.

Temperature further narrows acceptable options. Elevated process temperatures, localized hot zones, and thermal cycling can increase evaporation rates and accelerate degradation. A lubricant may deliver excellent wear protection but excessive volatility in vacuum.

Another may offer low outgassing but insufficient durability for high load motion. Selection should begin with the dominant failure mechanism in the specific component: molecular contamination, particulate generation, film instability, material incompatibility, wear, or thermal degradation.

Selection should start with component-level risk

Selection factors become repetitive when vacuum level, temperature, motion, load, and compatibility are treated as a checklist. They are more useful when tied directly to the risk they control. Vacuum level points to volatility and condensable molecular contamination.

Temperature affects evaporation, oxidation stability, and degradation. Motion profile influences film stability, oil migration, and particle generation. Load affects wear, surface fatigue, and protective film thickness. Material compatibility determines whether elastomers, plastics, coatings, or metals will swell, harden, crack, corrode, or otherwise change when exposed to the lubricant.

The more useful question is which performance profile reduces the most relevant failure pathways for that location in the tool. A seal interface may place material compatibility and low swelling tendency ahead of extreme wear life.

A linear guide in a vacuum tool may prioritize ultra-low volatility and controlled oil separation. A bearing near a thermal process may need high temperature stability, particulate cleanliness, and film durability under load. A robotic handling system may depend on consistent torque and long-term film stability through repeated cycling.

This is why lubrication should enter the engineering process before component geometry, materials, access points, and maintenance intervals are fully locked. Late selection can force a compromise between mechanical reliability and contamination control. Early selection gives engineers a better chance to align lubricant chemistry, component materials, tool architecture, and maintenance strategy before those conflicts become expensive to resolve.

It also helps fabs avoid both under-specifying and over-specifying. Some applications may need the most chemically robust lubricant available because the failure consequence is severe. Other subsystems may be able to meet the technical requirement with a lower volatility, non-fluorinated, or more compliance-aligned option. The point is to define the contamination sensitivity and mechanical requirement at the component level instead of applying a single lubricant assumption across the tool.



Chemistries such as perfluoropolyethers, multiply alkylated cyclopentanes, and polyphenyl ethers may be considered for demanding vacuum or high-temperature environments because they can provide low volatility and thermal stability. No chemistry is universally suitable, however.

Validation data should be structured, recognized, and application specific

Data sheets are useful, but semiconductor applications need validation that connects laboratory screening to real tool behavior. Recognized test methods provide a common basis for comparison and help engineers move beyond broad claims. ASTM E595 is commonly referenced for outgassing under vacuum and thermal exposure. Its Total Mass Loss measurement indicates the percentage of material lost after exposure, while Collected Volatile Condensable Material indicates the portion capable of condensing on nearby surfaces. For optics, sensors, metrology systems, wafers, and chamber surfaces, CVCM can be especially relevant because it points to redeposition risk rather than mass loss alone.

Vapor pressure characterization can further support vacuum qualification. Because direct room-temperature measurement can be difficult for high-molecular-weight lubricants, engineers often rely on elevated-temperature data and extrapolation to understand long-term volatility behavior under operating conditions. That information should be read alongside outgassing data, not as a replacement for it, because vapor pressure helps describe the tendency to enter the gas phase while CVCM speaks more directly to the risk of condensable deposition.

Particulate cleanliness should be evaluated with the same specificity. Standards such as IEST-STD-CC1246 classify contamination levels by allowable particle size and count, giving engineers a way to specify and compare lubricant cleanliness for sensitive environments. Particles do not only threaten wafer surfaces. They can disrupt precision motion, damage bearing surfaces, affect optical assemblies, and create secondary wear debris inside the tool. A lubricant may appear acceptable based on bulk properties while still carrying particulate contamination that is unacceptable for a high-precision subsystem.

Standardized screening is still only the first filter. Application-specific testing should evaluate dynamic behavior under relevant motion, load, temperature, and vacuum conditions. Motion rigs, thermal cycling, vacuum exposure, material compatibility testing, and controlled pilot trials can reveal risks that static tests do not fully capture, including oil migration, film breakdown, particle generation under shear, seal interaction, and changes in service interval.

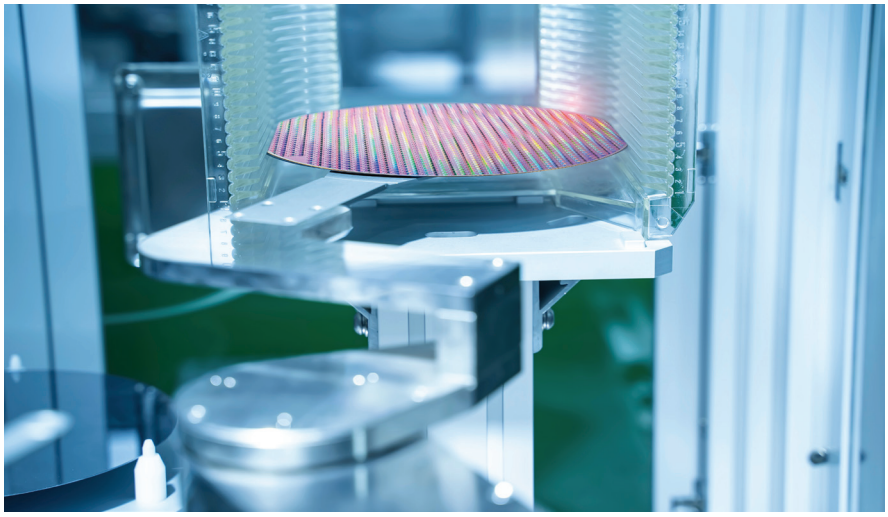
Production validation should be limited and intentional because semiconductor downtime is expensive. Candidates that perform well in earlier screening can be introduced into controlled tool sets while teams monitor particle counts, component wear, service intervals, surface cleanliness, torque behavior, uptime, and yield trends. The goal is not to prove that a lubricant performs well in general. The goal is to confirm that it can maintain cleanliness, film stability, wear protection, and process reliability under the actual conditions of the tool.

Supplier collaboration can make qualification more efficient when it is grounded in technical tradeoffs rather than product preference. In semiconductor tools, lubricant selection often requires engineers to weigh requirements that do not always align with each other: low outgassing, particulate cleanliness, film stability under cyclic motion, high temperature performance, wear protection, material compatibility, service life, and regulatory direction.

FUCHS brings value to this stage through its broad portfolio of semiconductor-relevant lubricant chemistries, technical application support, and ability to develop custom solutions when standard products do not fully match the operating environment. That breadth allows engineers to compare fluorinated and non-fluorinated options against the same performance criteria rather than defaulting to a familiar chemistry or over-specifying for margin. The objective is to determine whether a lubricant can maintain the specific balance of cleanliness, film stability, molecular control, and mechanical protection required by the tool—not only at installation, but over sustained vacuum, motion, and thermal exposure.

Regulatory pressure adds another chemistry constraint

PFAS and PFOA scrutiny has added complexity to semiconductor lubricant



selection. Fluorinated chemistries, including PFPE oils and PTFE thickeners, are discussed more closely because of concerns related to persistence and exposure. At the same time, the properties that drive regulatory attention are also part of what has made these chemistries valuable in demanding semiconductor tools.

Low vapor pressure, chemical inertness, temperature stability, and broad compatibility can be difficult to replace in aggressive vacuum, high temperature, or contamination sensitive applications.

The answer is not a one-for-one substitution exercise. Some applications may continue to justify fluorinated lubricants because alternatives cannot yet meet the full combination of vacuum performance, thermal stability, wear protection, cleanliness, and compatibility. Other applications may be able to transition to non-fluorinated synthetic chemistries, including certain MACs, PPEs, advanced esters, or specialty hydrocarbon systems, without compromising the functional requirement. The decision should be based on structured comparison against the same technical criteria, not on habit or chemistry category alone.

That comparison also protects the fab from creating a new operational risk while solving a regulatory one. A chemistry change that reduces PFAS exposure but increases volatility, seal degradation, or particulate generation may undermine the original objective of contamination control. Better decisions come from comparing fluorinated and non-fluorinated candidates against the same operating conditions, including vacuum performance, thermal stability, wear life, material compatibility, service interval, and long-term compliance alignment.

Better lubricant decisions protect yield before problems appear

As semiconductor tools become more complex and contamination tolerances become tighter, lubricant selection deserves the same engineering discipline applied to other process critical decisions. Lubricant related problems often remain invisible until condensable material has reached a sensitive surface, particles have affected motion repeatability, or premature wear has started to influence uptime. By that point, the lubricant has already moved from a component level decision to a process level problem.

Earlier evaluation changes that trajectory. When engineers identify likely failure mechanisms, validate chemistry through recognized and application specific testing, and compare materials against the actual operating environment, they reduce the chance that lubrication becomes an unexplained source of variability.

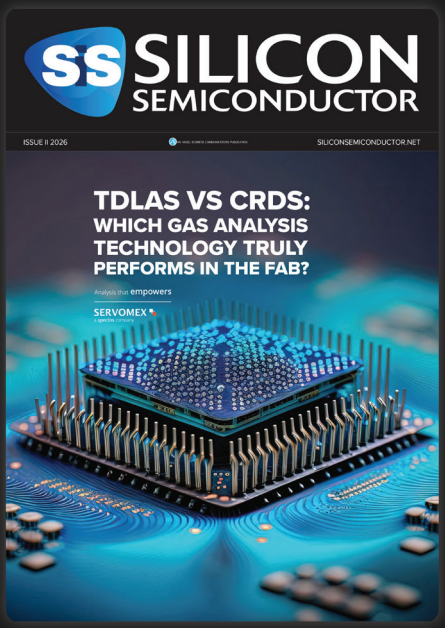
In a fab, precision is protected by thousands of decisions that keep the process stable. Lubricant selection is one of the smaller decisions on the drawing, but when it is made correctly, it helps protect yield before the problem ever appears.

SILICON SEMICONDUCTOR

BOOK YOUR REPRINT TODAY!

A reprint from **SiS Magazine** amplifies your editorial exposure and provides authoritative third-party validation. It's a powerful tool for **sales and marketing**, helping you showcase innovation with independent credibility. Use it across **events, pitches, newsletters, and social channels** to reinforce trust, extend visibility, and strengthen your position in the silicon semiconductor industry.





Contact: **Shehzad Munshi**
shehzad.munshi@angelbc.com



AI-designed silicon: Redefining the route from workload to accelerator

The development of specialised AI hardware has traditionally been a lengthy, sequential process. Architecture is defined, hardware is designed, verification follows, software is developed around the resulting implementation, and successive iterations are tested before a chip can eventually reach production. Architect Labs is proposing a substantially different model, in which artificial intelligence is used to connect these stages into a single optimisation process.



THE COMPANY has unveiled Redwood, an AI accelerator that it says was designed, verified and deployed on an FPGA in less than two weeks, starting from a specification written by two human architects. According to Architect Labs, its AI system generated the accelerator's performance model, RTL, verification environments, formal proofs, firmware, drivers and compute kernels without human intervention below the specification level. The work was also undertaken without using pre-existing accelerator IP.

The significance of Redwood therefore extends beyond the specifications of one accelerator. It represents an experiment in whether AI can automate not just individual stages of semiconductor development, but the interactions between architecture, hardware and software that determine whether a processor is well matched to a particular workload.

That distinction is important as AI moves into increasingly diverse applications. The computational requirements of large language models, computer vision, robotics and other forms of Physical AI are not necessarily identical. A processor designed around a specific combination of model architecture, latency, memory requirements and power constraints can potentially make different trade-offs from a general-purpose accelerator.

From sequential design to optimisation loop

Architect Labs argues that the conventional chip-development model creates a particular problem for AI hardware. Semiconductor development can take years, while AI models and workloads can change much more rapidly. By the time an architectural decision has passed through RTL development, verification and implementation, the workload for which the hardware was optimised may have evolved considerably.

The company's alternative is based around its Architect Labs Platform, or ALP an in-house platform for end-to-end chip design. Rather than treating architecture, RTL, verification and software as largely separate activities, the platform uses a high-level design specification as the common source of truth. Architectural alternatives can then be explored across hardware and software, with functional, area, performance, timing and power considerations used to evaluate different implementations.
This is potentially significant because hardware/software co-design is

not a new concept. What changes in the proposed model is the scale and speed at which alternatives can be investigated. If AI systems can generate, test and discard large numbers of candidate implementations automatically, architectural exploration becomes less constrained by the number of engineers who can work on a problem simultaneously.

For Redwood, Architect Labs says that each architectural iteration could be regenerated, reverified and redeployed to an FPGA within 48 hours.

The company also reports that every block through the SoC level achieved at least 95% code and functional coverage using commercial electronic design automation tools, its own formal engine and hardware-in-the-loop validation. These are company-reported results rather than independently established industry benchmarks.

The approach also changes the relationship between software and hardware. Rather than waiting for hardware to exist before optimising software, ALP is intended to develop firmware, kernels and performance models alongside the hardware architecture. This allows software implications to influence architectural decisions before those decisions become embedded in RTL.

Architecture for low-power inference

Redwood itself has been designed around single-batch, low-power and low-latency inference workloads. Its architecture uses a mesh of tiles connected through a custom on-chip network. Each tile combines a RISC-V control processor with dedicated compute resources for transformer inference, including INT8 matrix multiplication engines and a vector engine incorporating SIMD processing, transpose functions and floating-point activation units.

Each tile also contains 512KB of local memory. Rather than implementing every operation as dedicated hardware, the architecture attempts to reuse resources where appropriate. The company's implementation of softmax, for example, uses an emulated algorithm based on FlashAttention-4 and the existing SIMD resources rather than introducing separate hardware for the operation.

This type of specialisation illustrates one of the potential advantages of AI-assisted hardware exploration. Conventional accelerator design necessarily involves decisions about where to invest silicon area, memory and power. Dedicated hardware can improve performance for a particular operation but consumes resources that cannot easily be redirected. Software implementation provides flexibility but can impose additional computational overhead.

An automated search process could, in principle, examine more combinations of these trade-offs than a human design team could practically evaluate. Architect Labs says its system explored the performance, area, timing and coverage space of the Redwood SIMD engine over multiple days, generating and verifying substantially different candidate microarchitectures.

The important question is not whether AI can generate RTL. It is whether automated exploration can consistently identify architectures that provide meaningful advantages once all the constraints of a real chip are considered.

From FPGA demonstration to silicon

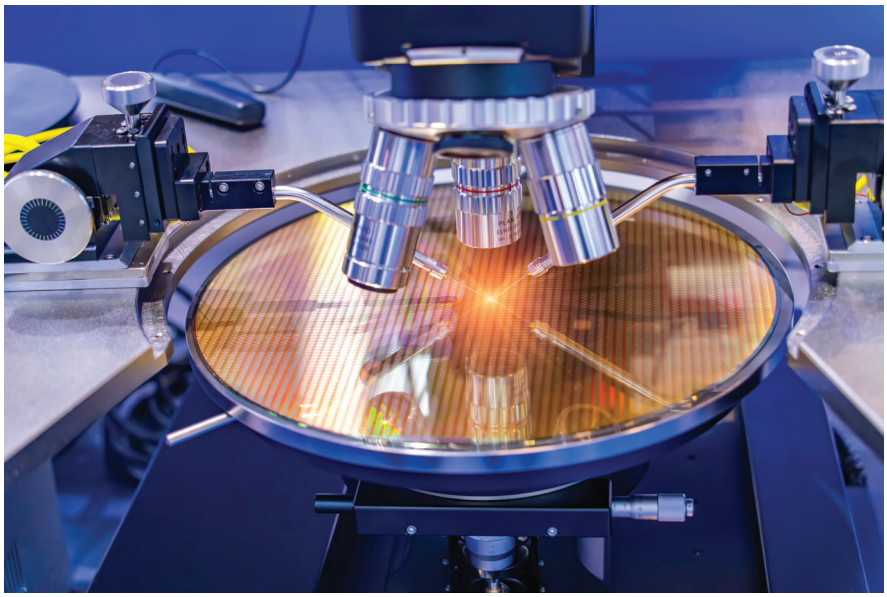
That question remains open because Redwood has not yet been demonstrated as a manufactured ASIC. The currently reported hardware results come from Redwood Nano, a scaled-down two-by-two tile implementation synthesised on an AMD Versal VPK180 FPGA and operating at 250MHz. It has been used to run Qwen3-0.6B end to end, including communication of the prompt to the accelerator and streaming generated tokens back to the host.

Architect Labs reports a throughput of 12.1 tokens per second for this configuration. It also projects that an ASIC implemented at the same process technology class as the NVIDIA Jetson Orin Nano could deliver a projected 3.4-times improvement in performance per watt compared with a measured Jetson baseline running the same model. The company further projects an order-of-magnitude improvement in area efficiency.

These figures should be distinguished from measured silicon performance. The company itself acknowledges that manufactured silicon represents the final validation of the design and says it is continuing development towards full GDSII generation and a TSMC tapeout.

That distinction is particularly important in evaluating an AI-generated chip. FPGA implementations, synthesis results and architectural projections cannot capture every characteristic of a production device. Physical design, process variation, clock distribution, thermal behaviour, memory performance, manufacturing yield and post-silicon validation can all influence the final result.

The transition from an FPGA



demonstration to a manufacturable chip is therefore one of the most important tests of the approach.

Verification becomes part of the experiment

Verification is arguably the most consequential part of the company's proposition. Generating RTL quickly is of limited value if the resulting hardware cannot be demonstrated to be correct.

Architect Labs says its system automatically generated testbenches, test cases, formal verification artefacts and simulations from the same design intent used to create the hardware. It combined UVM-based verification with formal methods and a proprietary formal engine.

According to the company, the first RTL version moved from simulation to the FPGA platform without identified bugs. The underlying concept is that verification capacity becomes linked to available computing resources rather than simply engineering headcount.

The company's AI agents can run large numbers of simulations and optimisation cycles concurrently, allowing alternative implementations to be assessed continuously. That potentially addresses one of the central bottlenecks in AI-assisted chip design. Increasing the number of generated designs is only useful if verification can scale with it. Otherwise, automated generation simply transfers the bottleneck from design engineering to design validation.

A hardware feedback loop

Perhaps the most unconventional aspect of Redwood is what Architect Labs describes as recursive self-improvement.

The company deployed Qwen3 on Redwood and exposed the model through an inference endpoint within its AI development system. It says the deployed model subsequently identified timing and kernel optimisations that could be incorporated into the accelerator's development.

This should not be confused with an autonomous system independently redesigning and manufacturing successive generations of silicon. At this stage, the demonstration is considerably narrower: an AI model running on the accelerator was used to identify optimisation opportunities for the hardware and software environment.

Nevertheless, it points towards an intriguing feedback mechanism. AI is being used to design the hardware; that hardware runs an AI workload; observations from that workload are then fed back into the development process.

If such a loop can eventually be extended through physical implementation, tapeout and post-silicon measurement, the potential implications are significant. Hardware development could become increasingly workload-specific, with future accelerators co-evolving with the models they are intended to execute.

Implications for Physical AI

This possibility is particularly relevant to Physical AI, where computational requirements extend beyond conventional cloud inference. Robots, autonomous machines and intelligent industrial systems need to combine perception, inference and physical action under constraints involving latency, power consumption, thermal management and physical size.

A general-purpose accelerator may offer flexibility across many workloads, but an autonomous machine could benefit from hardware optimised around the particular models and operating conditions it uses. Local inference can also reduce dependence on remote computing infrastructure where latency, connectivity or power consumption make cloud processing impractical.

Redwood is not itself presented as a dedicated Physical AI processor. Its architecture is described more broadly as targeting low-power, low-latency inference. However, the development methodology could become relevant to embodied systems if increasingly specialised accelerators are required at the edge.

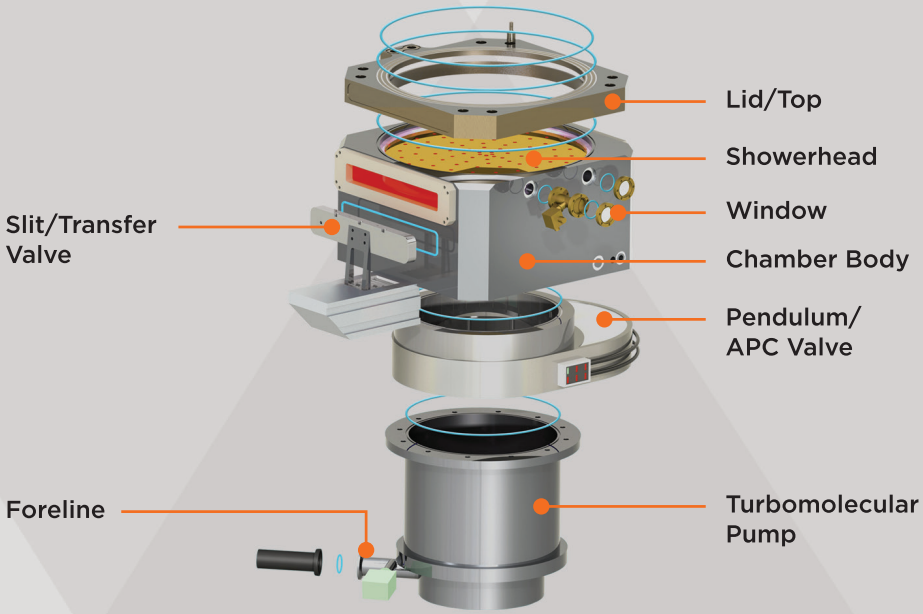
The broader proposition is therefore less about replacing conventional chip designers than about changing the rate at which specialised hardware can be explored. If AI can shorten the design cycle sufficiently, hardware could potentially follow rapidly changing workloads rather than requiring architects to predict those workloads several years ahead.

That remains a hypothesis to be tested. For now, Redwood provides an early demonstration of a different semiconductor-development model: one in which AI participates across architecture, hardware, verification and software rather than being confined to an individual stage of the engineering process. Its FPGA implementation provides tangible evidence that parts of this workflow can already be automated, while the projected ASIC results and planned tapeout represent the next, considerably more demanding stage of validation.

If that transition succeeds, the more consequential innovation may not be Redwood's particular architecture. It may be the possibility of making the journey from workload specification to specialised silicon substantially shorter, more iterative and more computationally driven.



Custom Design and Manufacturing Services for Applications Where Failure Isn't an Option



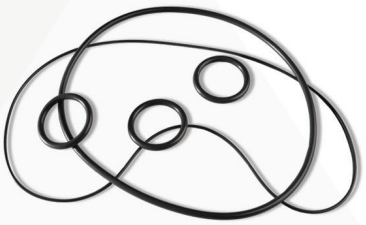
Seals and Technology That Keep Your Tools Running



Chemraz® G38
Minimal Contamination in Dry Plasma Etching Processes



Chemraz® XPE
Superior O2 Plasma Resistance to Protect Critical Components



Chemraz® XCD
Outstanding Thermal Resistance and Cleanliness



Avalon® 56 HP
High-purity plastic components for wet applications



ONX® 600
Fluoropolymer Composite with Excellent Acid Resistance at High Temperatures



Fusion® F07
High-performance FKM for Subfab applications

WHEN IT CAN'T FAIL™

Meet Greene Tweed at SEMICON
SEMICON West | October 13 - 15
Booth # 733

SEMICON Europa | November 10 - 13
NMI Pavilion Booth # B1121



Rethinking CMP for advanced packaging

How flexible, data-driven and resource-efficient platforms can support chiplets, 3D integration and panel-level packaging.

BY ACCRETECH (EUROPE).

AS THE semiconductor industry enters an era defined less by traditional transistor scaling and more by heterogeneous integration, advanced packaging has become one of the most important innovation battlegrounds. Chiplets, 2.5D and 3D integration, fan-out architectures, and the emergence of panel-level packaging (PLP) are enabling new levels of performance, functionality, and cost efficiency. Yet these packaging innovations bring significant manufacturing challenges, particularly for one critical process: Chemical Mechanical Polishing (CMP).

For decades, CMP was regarded as a mature, highly predictable process used primarily to achieve the planar surfaces required in semiconductor fabrication. Today, however, the technology finds itself at the centre of a rapidly changing production landscape. New materials, diverse substrate sizes, increasingly complex package architectures, and mounting sustainability expectations are forcing equipment suppliers and semiconductor manufacturers to rethink what CMP systems must deliver.

The question facing the industry is no longer whether CMP can provide high-precision planarisation. Instead, it is whether modern CMP platforms can evolve quickly enough to support the demands of advanced packaging while simultaneously meeting increasingly stringent environmental and operational requirements. The answer lies in a new generation of flexible, digitally connected, and sustainability-focused CMP systems.

Advanced packaging changes the rules

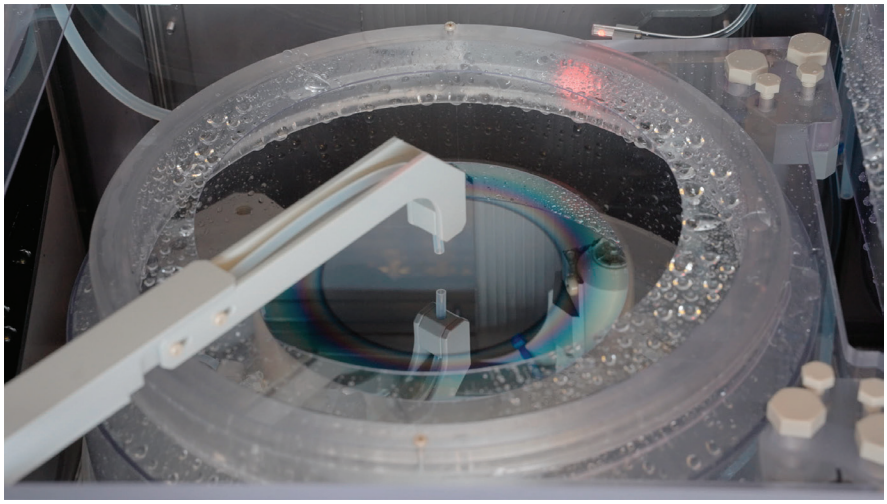
The semiconductor industry's long-standing reliance on Moore's Law has gradually given way to a more diverse approach to performance enhancement. Rather than simply shrinking transistors, manufacturers are increasingly combining multiple specialised dies into sophisticated package-level systems.

Chiplet-based architectures illustrate this trend clearly. Instead of developing increasingly large monolithic chips, designers can integrate separate compute, memory, power management, RF, photonic, and acceleration functions into a single package. This approach delivers advantages in performance, yield, design flexibility, and cost. However, such heterogeneous integration significantly increases manufacturing complexity. Surfaces must remain exceptionally flat to support reliable interconnections between disparate dies and substrates. Any topographic variation can negatively affect bonding quality, electrical performance, and long-term reliability.

CMP therefore becomes a process enabler rather than simply a polishing step. As packaging architectures become more sophisticated, the tolerance for process variation decreases dramatically.

At the same time, advanced packaging technologies are introducing a much broader range of substrates and geometries. Traditional production environments dominated by 300mm and 200mm silicon wafers are

increasingly giving way to mixed-format manufacturing. Facilities may now process 150mm, 100mm, or even smaller research-oriented wafers alongside conventional production substrates. Panel-level packaging expands this diversity even further by replacing circular wafers with large rectangular or square panels that offer improved manufacturing economics and area utilisation.
Such variation creates significant challenges for conventional CMP tools that were originally designed around standardised wafer formats. Equipment flexibility is becoming as important as polishing performance itself.



Material diversity presents new challenges

Just as substrate formats are changing, the range of materials entering CMP workflows is also expanding dramatically.

Historically, CMP equipment was optimised primarily for silicon and silicon dioxide processing. Today's advanced packaging applications involve a far wider material spectrum. Wide-bandgap semiconductors such as silicon carbide (SiC) and gallium nitride (GaN) are increasingly important in power electronics and high-frequency applications. These materials possess mechanical characteristics very different from silicon, requiring specialised process conditions to achieve the desired removal rates and surface quality.

Meanwhile, advanced packaging introduces softer and often chemically sensitive materials, including polyimides, build-up dielectrics, packaging resins, and other polymer-based layers. These materials demand entirely different polishing strategies, pressure profiles, slurry chemistries, and endpoint control techniques.

For manufacturers, this creates a significant operational challenge. The ability to switch rapidly between vastly different material classes without extensive cleaning procedures, contamination risks, or lengthy requalification periods has become a major competitive advantage.

Modern CMP equipment must therefore provide both process flexibility and process stability. Achieving one without the other is no longer sufficient.

The rise of high-mix manufacturing

Another consequence of advanced packaging is the emergence of increasingly diverse production environments.

Unlike traditional high-volume semiconductor manufacturing, many advanced packaging initiatives involve shorter product lifecycles, smaller batch sizes, and accelerated development schedules.

Research institutes, specialist foundries, start-ups, and established semiconductor companies alike are pursuing rapid innovation cycles to capture emerging opportunities.

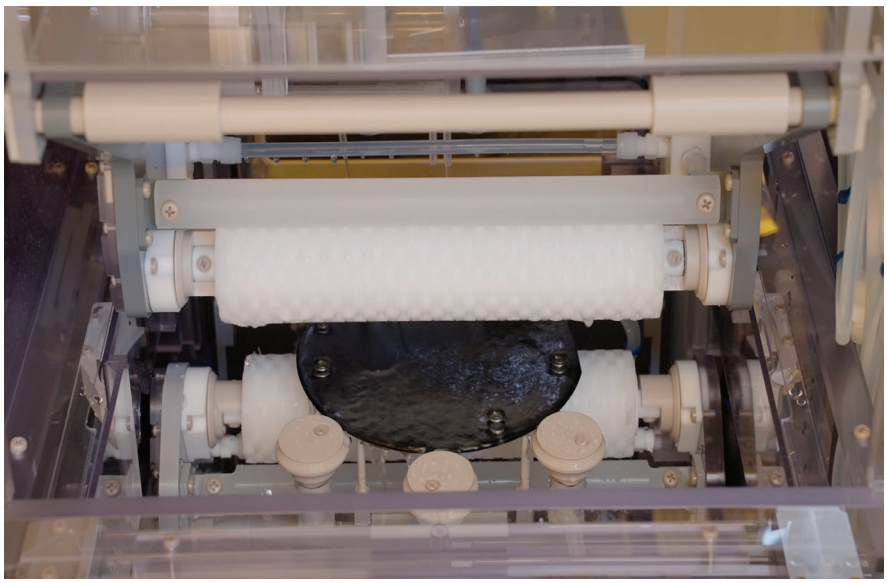
These environments require CMP platforms capable of supporting frequent recipe development and process experimentation. Engineers need the ability to evaluate new materials, optimise novel integration schemes, and quickly transition from proof-of-concept work to limited-volume production.

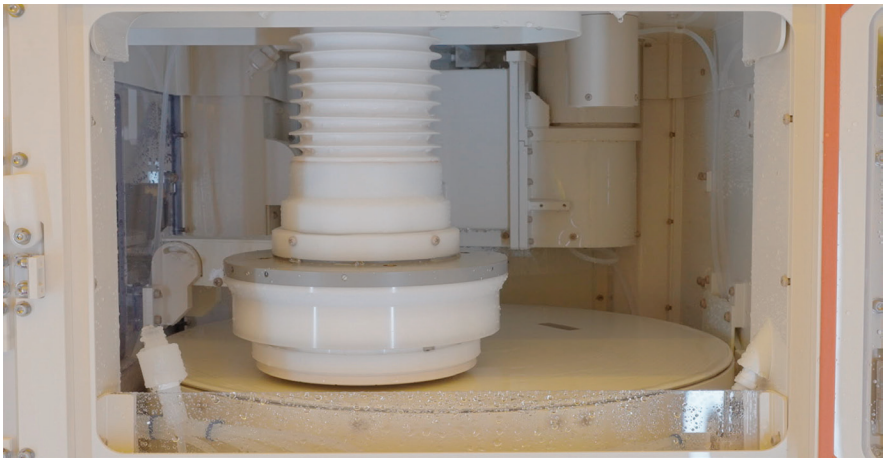
According to the white paper, flexibility in substrate handling, process switching, and throughput optimisation is rapidly becoming a baseline requirement rather than a premium capability.

The most successful CMP platforms of the future will be those that can accommodate this high-mix, high-variation manufacturing environment without sacrificing repeatability or yield.

Sustainability moves to centre stage

While advanced packaging is reshaping technical requirements, sustainability concerns are simultaneously transforming how fabs evaluate manufacturing processes.





CMP has long been recognised as one of the most resource-intensive steps in semiconductor production. Its combination of abrasive slurries, extensive rinsing operations, and demanding cleanliness requirements creates significant environmental impacts.

The white paper estimates that CMP can account for approximately 40% of total ultrapure water (UPW) consumption within semiconductor manufacturing operations.

Additionally, polishing a single 200mm wafer can generate between 30 and 50 litres of slurry waste containing abrasive particles and chemicals that are difficult to recycle.

Energy consumption is also substantial, with CMP processes contributing approximately 5-6% of overall fab electricity usage during operation. Life-cycle assessment studies have highlighted several major contributors to CMP's environmental footprint. The production and disposal of slurry chemicals, particularly formulations containing ingredients such as benzotriazole and hydrogen peroxide, represent significant environmental burdens. The energy needed for polishing, pad conditioning, and ultrapure water generation further increases process impacts.

As governments tighten environmental regulations and semiconductor companies pursue ambitious ESG targets, reducing these impacts has become a business priority. The industry's sustainability expectations are evolving rapidly. Throughput and yield remain critical metrics, but water consumption, energy efficiency, waste generation, and carbon emissions are increasingly influencing purchasing decisions and process development strategies.

Designing CMP for a more sustainable future

Meeting sustainability objectives does not simply require operational improvements. Instead, environmental performance is becoming an integral component of tool design.

One of the most significant developments involves wastewater management. New CMP systems increasingly incorporate on-tool waste stream separation technology, enabling fabs to segregate effluents according to chemical composition and contamination levels.

This approach simplifies downstream treatment and creates opportunities for partial water recycling. In some cases, rinse water generated during relatively clean oxide polishing processes may be suitable for recovery and reuse within fab water systems.

Slurry delivery technology is also improving. Modern injection systems distribute slurry more efficiently across wafer surfaces, increasing utilisation rates while reducing overall chemical consumption. Better distribution uniformity not only cuts waste but can also improve process consistency and reduce defect generation.

Pad conditioning systems are undergoing similar optimisation. Real-time monitoring allows tools to adjust conditioning pressure and frequency dynamically, extending pad lifetime while maintaining stable material removal rates. Reducing pad consumption delivers both operational and sustainability benefits.

Energy efficiency has emerged as another major development area. Advanced motor technologies, intelligent standby modes, and thermal recovery systems are helping reduce overall tool power requirements. Some equipment manufacturers have even expanded sustainability efforts beyond tool operation, incorporating renewable energy into manufacturing facilities and optimising logistics to minimise transportation-related emissions.

These initiatives reflect a broader industry shift: sustainability is increasingly being engineered into CMP platforms rather than treated as a separate compliance activity.

Flexibility through modular design

To address both technical and environmental challenges, equipment architecture is becoming increasingly modular.

The white paper highlights the importance of platforms capable of supporting multiple substrate formats without requiring entirely separate tool investments. Modern systems can often transition between 6-inch, 4-inch, and even 2-inch wafers within a single platform, with changeovers measured in hours rather than days.

Such flexibility delivers obvious financial advantages, reducing capital expenditure and increasing tool utilisation. However, the environmental benefits are equally important.

Avoiding duplicate equipment reduces manufacturing footprints and resource consumption. Faster transitions between process types minimise downtime and reduce the need for extensive cleaning cycles, conserving water and chemicals.

Modular process chamber designs represent another important innovation. By dedicating different chambers to specific material sets or process requirements, manufacturers can minimise cross-contamination risks while reducing cleaning frequency. This enables more efficient operation across diverse process flows.

Similarly, advancements in pad handling allow faster replacement and calibration while maintaining process consistency. Some systems

now support programmable pad zones capable of polishing different substrates using tailored process recipes on the same platform. Together, these innovations significantly improve equipment adaptability while supporting sustainability objectives.

Digitalisation transforms process control

Perhaps the most transformative development in next-generation CMP is the integration of automation and digital intelligence.

Historically, many process adjustments relied heavily on operator experience and offline analysis. Today's manufacturing environments increasingly demand real-time feedback, predictive control, and data-driven decision-making.

Advanced CMP platforms are therefore incorporating automated wafer handling, endpoint detection, and precision pressure control systems designed to eliminate variability and reduce operator intervention. Beyond the tool itself, integration with Manufacturing Execution Systems (MES) and Enterprise Resource Planning (ERP) platforms enables smarter production scheduling.

For example, grouping wafers with similar process requirements can minimise chamber transitions and reduce the number of rinse cycles required between runs. Such optimisation lowers water consumption, decreases downtime, and improves overall equipment effectiveness. Predictive maintenance capabilities are becoming equally important. By continuously monitoring polishing rates, pad wear, vibration signatures, and other process parameters, CMP tools can identify emerging issues before they result in scrap or unplanned maintenance events.

This not only improves yield but also reduces wasted materials and energy associated with failed processing runs. Digital traceability further strengthens sustainability efforts. Detailed logging of slurry use, energy consumption, water flow, and process conditions enables fabs to quantify environmental performance with increasing precision. Such capabilities support both regulatory compliance and internal ESG reporting requirements.

Looking towards AI-driven CMP

The evolution of CMP is unlikely to stop with today's automation technologies. Future systems are expected to leverage artificial intelligence, simulation-driven process development, and digital twin technologies to optimise performance and sustainability simultaneously. According to the white paper, AI-enabled control architectures and predictive diagnostics will become important tools for managing increasingly complex process environments.

Digital twins offer particularly compelling possibilities. By creating virtual representations of CMP processes, engineers can test process changes, evaluate new materials, and optimise recipes before committing actual wafers to production. This approach can dramatically accelerate development cycles while minimising waste, a critical advantage for advanced packaging applications characterised by rapid innovation and limited production volumes.

As these capabilities mature, sustainability metrics may become as tightly controlled as traditional manufacturing KPIs. Rather than measuring environmental performance retrospectively, fabs could establish resource consumption targets per wafer pass, covering parameters such as litres of ultrapure water, grams of slurry, and kilowatt-hours of energy consumed. CMP tools would then be expected to meet those targets by design.

CMP as a strategic technology

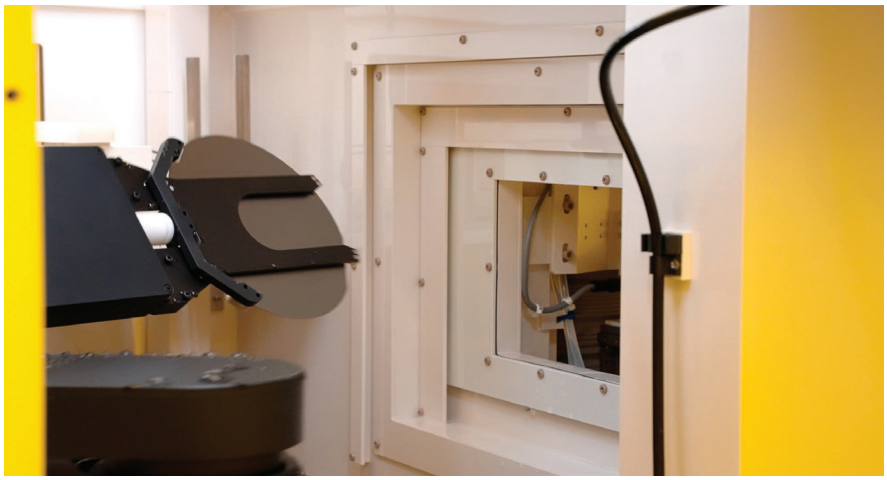
The semiconductor industry's packaging revolution is fundamentally changing perceptions of CMP. No longer viewed as a routine backend process, CMP is emerging as a strategic technology that enables the successful implementation of advanced integration schemes. Its ability to deliver precise planarisation across increasingly diverse materials and package structures directly influences product performance, manufacturability, and cost.

At the same time, the process sits at the intersection of two of the industry's most important priorities: advanced packaging and sustainability.

The next generation of CMP platforms must therefore achieve a delicate balance. They must be flexible enough to support chiplets, heterogeneous integration, PLP, wide-bandgap materials, and future packaging innovations. Yet they must also minimise water use, reduce chemical waste, lower energy consumption, and provide the data transparency demanded by regulators and corporate sustainability programmes.

The emerging consensus is that these goals are not mutually exclusive. In fact, the most adaptable systems are often the most efficient. Platforms designed to switch rapidly between materials, automate process optimisation, and integrate environmental management capabilities can simultaneously improve productivity and sustainability.

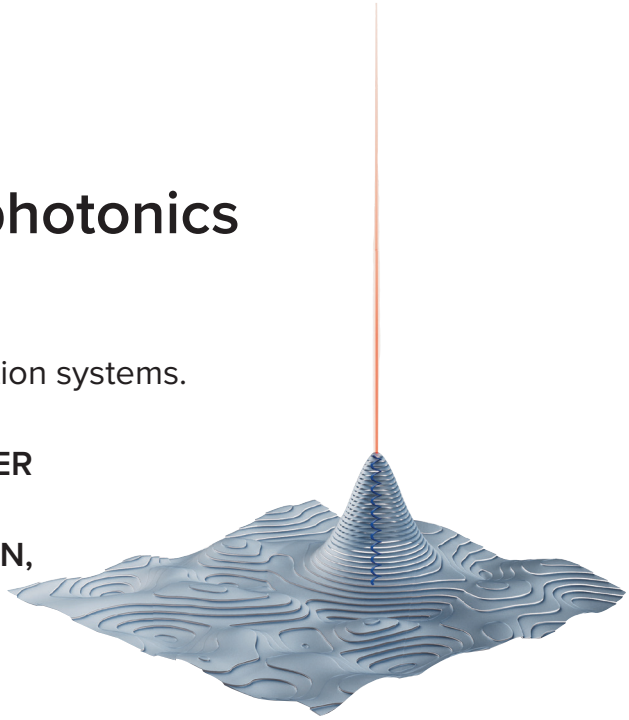
As advanced packaging becomes increasingly central to semiconductor innovation, CMP's role will only grow in importance. The tools that combine precision, flexibility, intelligence, and sustainability will help define the future of semiconductor manufacturing and determine how effectively the industry navigates its next major transition.



How to improve active photonics alignment performance

Latest advancements in algorithms and motion systems.

BY SCOTT JORDAN, BUSINESS DEVELOPER – PHOTONICS AND NANO AUTOMATION TECHNOLOGIES, AND STEFAN VORNDRAN, VICE PRESIDENT MARKETING – NORTH AMERICA, PHYSIK INSTRUMENTE L.P. (PI)

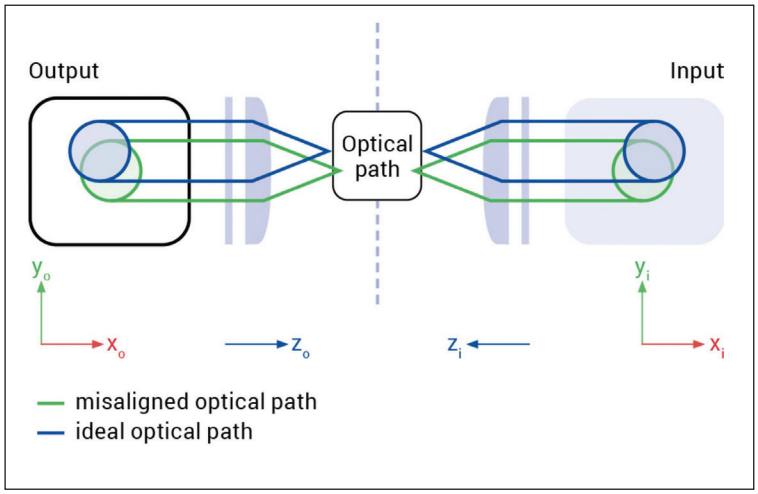


Why photonics applications are expanding

The world's appetite for data continues to grow at a steady pace, with new applications emerging every week. Even during the height of the COVID-19 pandemic in 2020, data consumption grew by more than 20 percent, making it vital for information to be efficiently routed and conveyed.

The photonics industry is growing alongside, as its technologies are found in many mainstream applications. These include various light detection and ranging (LiDAR) solutions and sensors in wearables and vehicle cameras, as well as functions in quantum computing, high speed cables, and even photonic computing where photons, rather than electrons, are used for logic. Silicon photonics (SiPh) is the technology behind these innovations, which combine optics and electronics on the same wafer to deliver significant performance gains in a smaller final product.

Based on the expanding and diversifying applications of SiPh, especially with the emergence of consumer applications, production volumes need to increase 1,000-fold in the next few years to keep up with the demand. This paper discusses the latest advancements in alignment automation that will help to make this possible.



➤ Figure 1: Testing and packaging today's photonic devices can be a huge challenge across multiple DOF. The alignment of multi-channel devices, such as fiber-optical arrays, used to be a slow, repetitive process before modern parallel algorithms were developed. (Image courtesy PI).

The difference between optics and electronics

All optical circuits first need to be tested, starting with the wafer and continuing through multiple stages of assembly. This requires much greater physical accuracy than probing electrical circuits, often two to three orders of magnitude higher, as well as additional degrees of freedom (DOF).

While probes can easily reach electrical test points, since pad sizes of approximately 30 microns are practical targets, SiPh structures need accuracy between 20 and 50 nanometers, calling for novel technologies with greater precision. This precision cannot be resolved with cameras or microscopes, and alignment must be 'active', meaning the positioning must be based on optimizing the optical power.

The challenge of upscaling SiPh production

The production of photonic integrated circuits involves particularly demanding nanopositioning

tasks, and vision- or fixturing-based approaches are inadequate, especially as variations exist between devices. The need for repetitive alignment for all manufacturing stages, from initial probing and grading on the wafer through final assembly and packaging, including multiple steps of placing and bonding functional elements interspersed with more testing, has long been regarded as an obstacle for scaling up production, as it accounts for up to 80 percent of the total costs. In fact, the traditionally high cost of active alignment methods has been one of the main drivers of the multi-decade search for passive alignment technologies. Nevertheless, precision active alignment is now possible, eliminating many of the associated costs and enabling a cascade of savings across the manufacturing workflow.

Passive or active alignment?

In non-photonics manufacturing, precision alignment can be achieved using either an active or passive strategy. An example of the latter is a dovetail joining two boards at right angles to each other, where the edges perfectly align. However, passive positioning is often not practical for applications requiring higher precision, such as the alignment of optical fibers to chips or other components, because the tolerances are too tight, which makes the manufacturing processes unreliable.

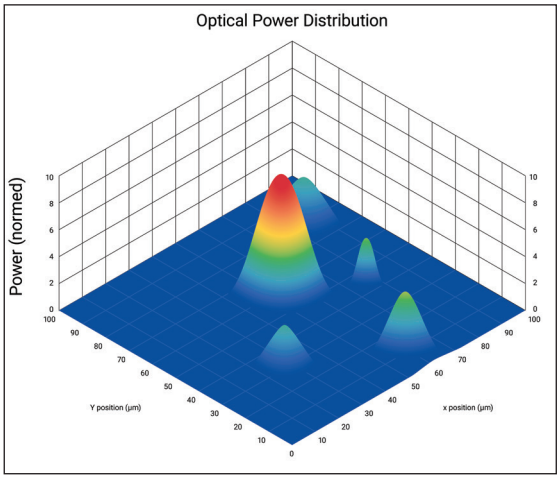
Furthermore, as the design becomes more technical, the machine tools and processes that can meet the required tolerances grow more expensive and, inevitably, the number of rejects rises, resulting in additional costs. Besides, there are some device tolerances that cannot be addressed with better fabrication techniques, such as fiber-core centration, and machining the world's finest V-groove is of little value when device-to-device variability exceeds the necessary alignment tolerances.

Nevertheless, clever passive alignment solutions have emerged that show great promise for specific applications, such as the Photonic-Plug (Teramount), but the need for active alignment is not going to disappear.

Comparatively, active alignment maximizes performance through the use of robots to automatically align devices and, given the cost savings, flexibility and speed described below, it will remain the preferred choice for photonic device manufacturing in the future. Active alignment works by optimizing coupling performance, and the latest innovation described in this paper performs this cost-effectively across numerous channels in multiple DOF, which is required as SiPh has the ability to print several circuits with multiple inputs and outputs on a single chip.

Active alignment: what is necessary?

Active alignment requires intelligent control electronics and high precision mechanics. It involves autonomously controlling the mutual position of the photonic device's components during the manufacturing process.



➤ Figure 2: Power distribution of an optical component shows multiple peaks. Modern algorithms and precision mechanics can help determine the main mode and find the exact peak quickly. (Image courtesy PI).

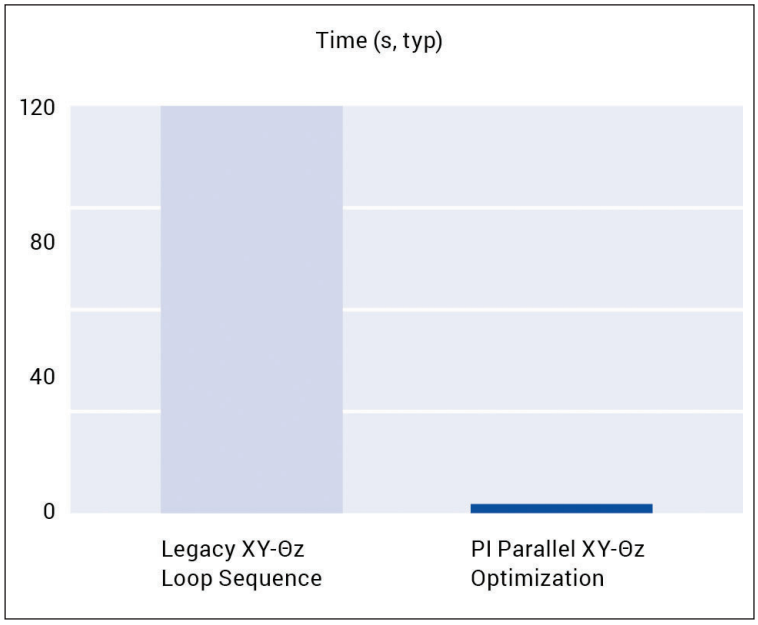
Traditionally, this has been a time-consuming process, because optimization of several channels, inputs and outputs in multiple DOF, required numerous iterations until the result was right. Using legacy approaches, this time will only increase as devices grow smaller and more complicated, increasing the cost further. For example, alignment in a multi-lens assembly is far more demanding than for a single element, as is a SiPh chip with multiple channels instead of the single fibers people were attaching in the late 1990s. Fortunately, breakthrough control algorithms can now perform all the necessary sub-alignments across multiple DOF, channels and operating systems in parallel, accomplishing the whole job in one quick step, while eliminating repetitive looping iterations and reducing alignment time by typically 99 percent.

Precision positioning of signal-carrying elements

The central challenge in active alignment is accurately positioning signal-carrying elements, including single fibers or fiber arrays, lenses, interposers or other chips, with respect to coupling points such as grating couplers or edge facets. These couplings can be into or out of passive and active photonic structures, for example, waveguides, vertical cavity surface emitting lasers, or photodiodes. Typically, alignment accuracy down to tens of nanometers is mandatory when coupling photonic components, to generate the highest possible optical power transmission and lowest attenuation.

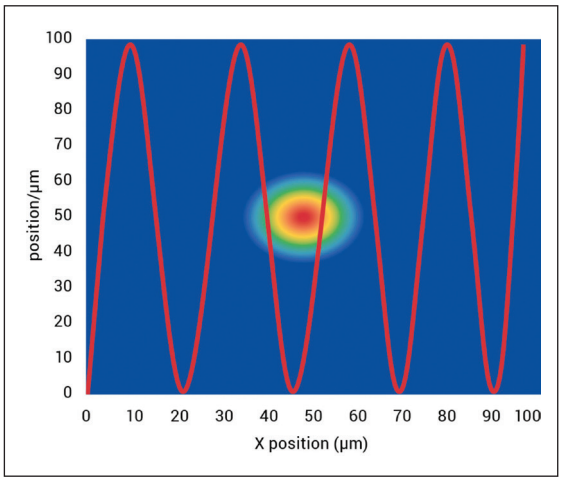
Precision automation: Finding the main peak

In almost all photonic component alignments, the signal, for example coupled optical power, rises to a global peak when optimally aligned and falls when out of position. This is generally a universal pattern, and often false lower or local peaks can be seen, which must not be selected for optimization. Finding that global peak, however, has traditionally been a

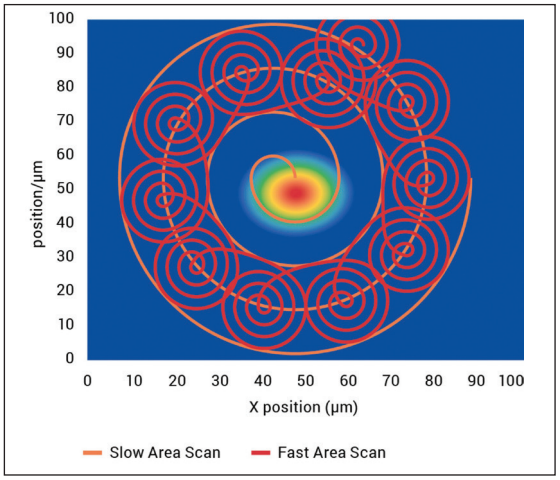


➤ Figure 3: Alignment speed is crucial in reducing production costs. (Image courtesy PI).

time-consuming task, especially when it depends on several DOF.



➤ Figure 4a: Sinusoidal area scan for detecting first light. (Image courtesy PI).



➤ Figure 4b: Spiral scan using a hexapod/piezo approach. A fine and coarse scan can be executed simultaneously. (Image courtesy PI).

Advanced algorithms and fast mechanics

A novel combination of advanced algorithms and rapid mechanisms, for example, the award-winning Fast Multichannel Photonics Alignment solutions (FMPA) from PI have largely accelerated throughput. FMPA simultaneously optimizes coupling between photonic elements across channels, inputs and outputs, and DOF, even when these independent variables influence each other, through novel functionalities that include rapid first light detection, a vibrationless area scan, and an innovative parallel gradient search that can perform real-time, multivariate tracking optimization.

Coupling in previous approaches had to be optimized in a sequential, looping fashion, constantly going back and forth and readjusting at each axis to gradually reach a consensus. This took a considerable amount of time, often several minutes, which had a significant impact on costs, efficiency and scalability.

Comparatively, FMPA can achieve all of these adjustments in just one single step with a simple set of commands, which can be performed in just seconds. Furthermore, optimization time is largely independent of the number of alignments being performed, since they are performed in parallel. This significant reduction in alignment time, typically two orders of magnitude, massively reduces costs, making processes such as probing SiPh wafers economically viable.

Active alignment: different processes

There are two fundamental types of alignment processes for photonic components: Area scans and gradient searches. Area scans determine the peak of a measured figure of merit, such as optical power, modulation transfer function (MTF) or modal purity, within a defined region, and can be used to precisely characterize the optical fingerprint of a component, as well as separate global and local maxima. Gradient searches provide fast final optimization and tracking, and FMPA's unique implementation can optimize one or more couplings in multiple DOF at once, as well as track them to mitigate drift processes, disturbances, etc. These processes are discussed in more detail below.

Area scans

Area scans, which scan an area to determine the approximate location of the highest coupling peak, are used for a variety of tasks, including:

- Detecting first light;
- Characterising (profiling) a coupling, which is important for process control;
- Localising the main mode of a coupling for subsequent optimization by a gradient search. This sequence forms a powerful, hybrid approach that helps to prevent locking onto a local maximum. The area scans built into FMPA include unique single frequency sinusoid and spiral scans. These are much faster than raster or

serpentine scans, as they are truly continuous and avoid settling the requirements of the stop-and-start motions used in traditional scans. Additionally, the frequency is selectable to avoid exciting structural resonances. It is also possible to select a constant velocity spiral scan, which allows the acquisition of data with constant density across the spiral.

Gradient searches

Gradient searches perform a small, approximately circular dither motion between devices to modulate the coupling. This variation of the figure of merit allows the local gradient of the coupling to be determined in real time. The controller automatically observes modulation and drives the alignment to reduce the modulation until it falls to zero, indicating full optimization.

From the observed modulation, the controller can automatically deduce the local gradient via a simple calculation (Figure 6). Note, that the gradient (Δ I) falls to zero when optimized. Any axes in an FMPA system can perform these types of alignments, subject to the physical capabilities of the axes.

Gradient searches are rapid and precise tools for transverse optimization, but they can also be used in other functions, including:

- In a single linear axis, which is ideal for localizing the beam waist in a lens coupling;
- In a gimbaling or swivelling fashion to optimize an angular orientation; or
- In a rotation about one channel axis of a multichannel device, to bring all elements of the device's arrayed input/output (I/O) channels into correspondence.

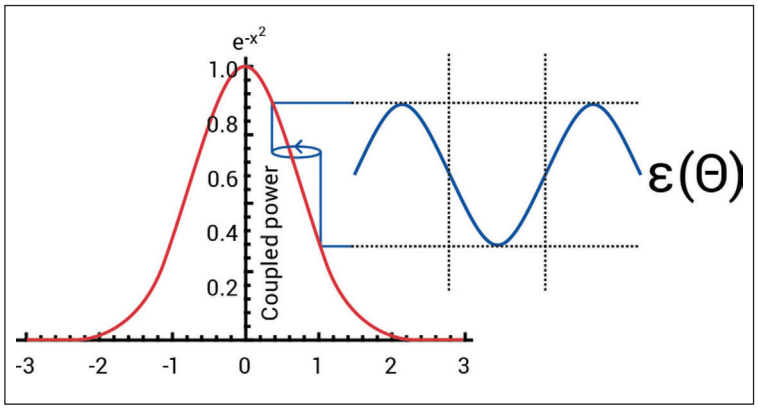
These algorithms are suitable for all kinds of optimizations, including bulk optic, cavity and pinhole alignments.

Once the main mode has been identified through an area scan, the algorithm will reach the peak quickly. The main peak is automatically fine-tuned in all DOF the application requires. All of this is achieved in parallel with one command by the user. With today's photonic devices, it cannot be assumed that the peak is symmetrical follows a Gaussian distribution, or has a circular cross-section. Furthermore, both edge and diffractive-coupler scenarios must be met.

Compensating for position changes

In packaging, photonic devices must be permanently bonded together and connected to fibers to be able to transmit signals, which is typically done with a UV-curing epoxy resin.

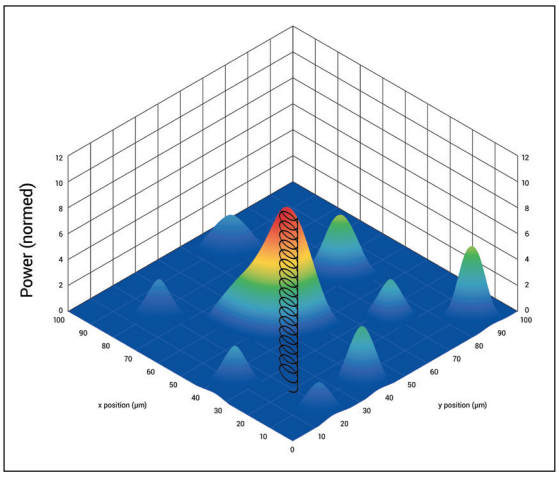
However, polymerization that occurs during the curing process increases stress and causes displacement of the components. For this reason, a tracking mode within the FMPA algorithm is useful



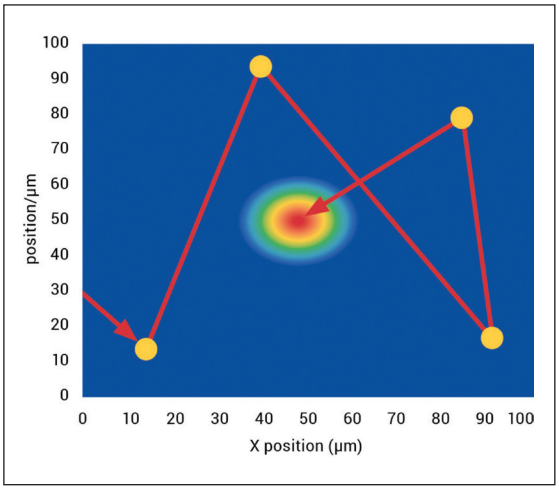
➤ Figure 5: Graphical depiction of gradient determination via a circular dither, which modulates the observed coupled power (or other parameters). The modulation phase with respect to the dither indicates the direction towards maximum while its amplitude falls to zero when optimized. (Image courtesy PI).

$$|\varepsilon(\Theta)| = \Delta I = (I_{min} - I_{max})/I_{min}$$

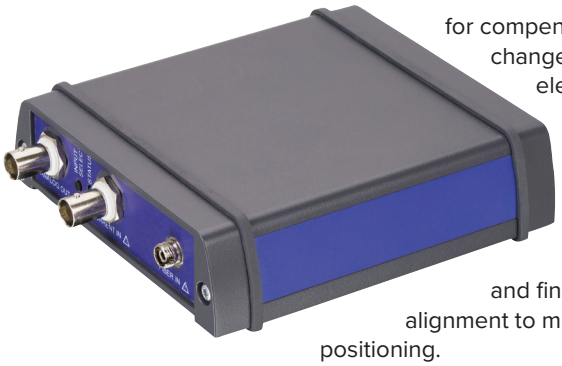
➤ Figure 6: The observed gradient can be expressed by the above equation and serves as a measure of alignment error. (Image courtesy PI).



➤ Figure 7: A graphical depiction of a digital gradient search (hill-climb algorithm). (Image courtesy PI).



➤ Figure 8: First light detection with PILightning. (Image courtesy PI).



for compensating positional changes of the optical elements in real time, in the early stages of polymerization, making the necessary adjustments, and fine-tuning the alignment to maintain optimum positioning.

➤ Figure 9: High speed optical alignments require optical power meters with high bandwidths. (Image courtesy PI).

Solving the first light problem

Before the aforementioned optimization processes can even start, an optical signal, above the noise level, needs to be detectable. This process is called first light detection. Finding first light has been a time-consuming procedure in all industrial photonics alignment applications, including wafer probing and device packaging. It is particularly arduous in devices with inputs and outputs where both sides must be aligned to achieve even a threshold level of coupling.

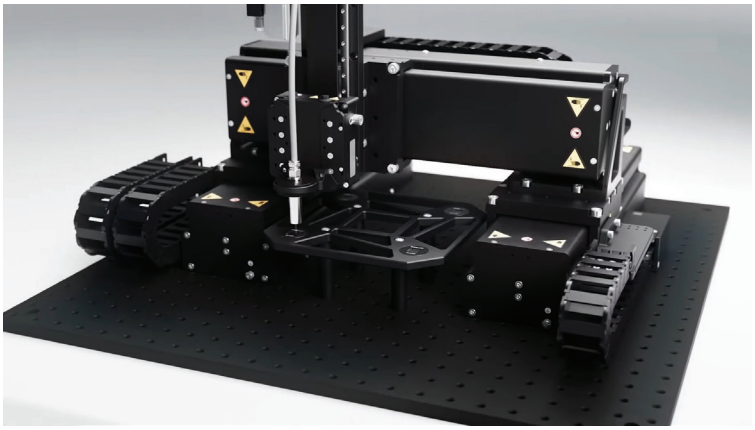
Traditional first light search algorithms

Signal finding is typically based on performing cyclic patterns such as Archimedean spirals or sinusoidal raster scans at the micron to submicron scale (see Figures 4a and 4b). In the case of large device-to-device variations or indeterminate fixturing, these repetitive, tightly spaced scans can take a significant amount of time to complete depending on factors such as the area to be searched, whether inputs and outputs need to be simultaneously aligned, and so on.

New first light capture process

A breakthrough has been achieved in the form of a novel, embedded search and alignment algorithm (patent pending), that promises to revolutionize the field. The algorithm, called PILightning, runs embedded on PI's advanced controllers. It enables highly dynamic mechanics, such as piezo scanners or direct-drive air bearing stages, to achieve significant economic gains in production over previous first light search algorithms. This new process is fully automated and virtually instantaneous, eliminating the need for extensive calibration or manual intervention.

➤ Figure 10: Mini Gantry. (Image courtesy PI).



PILightning is based on a new search method with integrated AI-based real-time executive function (see Figure 8). It dramatically reduces the time required to find first light in single-sided and double-sided couplings and in loopback (omega) waveguide configurations. Once first light is detected, the FMPA fast gradient search algorithm takes over, utilizing real-time feedback control to rapidly optimize the alignment in parallel across degrees of freedom and channels. Depending on the application, a tracking algorithm can also be activated to maintain maximum coupling efficiency – important, for example, in curing situations.

Orders of magnitude improvement

Tests have shown that PILightning typically reduces first light capture by one order of magnitude or more in single-sided alignment applications. Gains of more than two orders of magnitude are achieved in double-sided applications. The larger the search area and (as with the FMPA parallel optimization functionality) the more complex the alignment, the more significant the gain.

Mechanisms for different requirements

While the algorithms mentioned above apply to all kinds of mechanical alignment hardware, various demands, such as size, travel range, DOF, etc. on alignment mechanisms require different configurations and technologies.

Optical power meters: bandwidth and dynamic range rule

A highly sensitive power meter with wide dynamic range and high bandwidth is also crucial to support the optimal alignment of SiPh components. Typically, 20 kHz signal bandwidth and a dynamic range of six orders of magnitude with logarithmic output are leveraged to make full use of the FMPA algorithms.

Gantry platforms for large format applications

Processing densely packed nanoscale structures over large areas for various applications, including semiconductors and SiPh testing and manufacturing across circuit boards, trays, carriers and other large substrates, requires high precision, high throughput solutions delivered on reliable platforms. Novel gantry systems are suited perfectly for these applications, helping to significantly improve throughput and deliver repeatable and robust results in a compact form factor.

Some modern gantry systems, such as those provided by PI, use power-dense ironless linear motors to accommodate the high dynamics required in photonics applications. These efficient direct drive motors can have two high resolution linear encoders, with nanometer or sub-nanometer resolutions, on each gantry axis, which deliver repeatable positioning even in the most demanding

duty cycles. Mechanical or air bearings are available, with the latter providing multiple advantages, including a cleaner operation with no risk of contamination that could damage the delicate optical circuits. Controllers with and without built-in FMPA functionality are available.

Hexapod mechanisms: Parallel kinematics

Hexapods are extremely versatile motion and positioning systems, which are based on six actuators placed between two plates operating in parallel. They provide all six DOF in a very compact and stiff package, making them an obvious choice for photonic integrated circuit (PIC) test and packaging applications. These require at least four DOF of alignment automation with a flexible variable rotational pivot point, and often benefit from five or six axis optimizations. Additional power and flexibility come from mixing and matching different motion technologies. For example, the addition of a 3-axis piezo flexure scanner offers unmatched nanoscale resolution and accelerates the alignment process due to the low inertia of the piezo mechanism, which leads to a high scanning frequency to increase productivity and generate a tight, responsive lock-on.

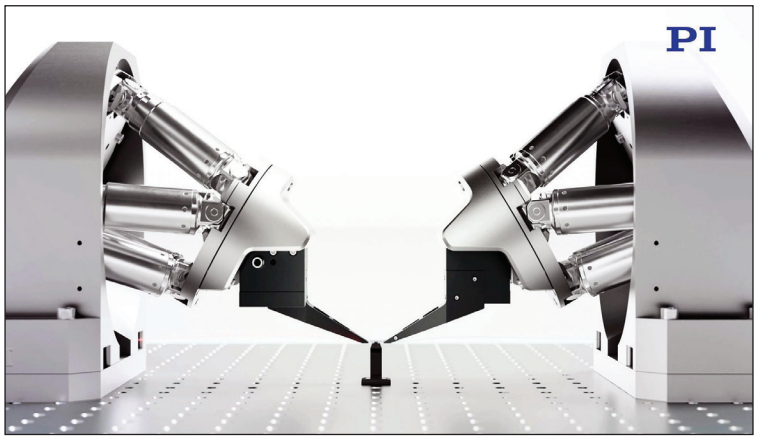
PI hexapod controllers also offer a programmable coordinate system, allowing them to be mounted at any angle while maintaining the desired XY plane. This also provides a flexible definable pivot point that allows the user to specify the center of rotation anywhere inside or outside the hexapod envelope. The definable pivot point is invaluable when rotation around a fiber tip or beam waist is required.

Modular alignment systems: Serial kinematics

Modular alignment systems that consist of a combination of individual stages, known as serial kinematics, are guided by one of two designs: mechanical or air bearings. Air bearings are preferred as they provide precise, high dynamic motion that is free of friction, wear, and maintenance. Furthermore, air bearings do not generate particles as they neither contain lubricants nor are subject to abrasion caused by friction.

Mechanical bearing stages

Mechanical bearing stages are cost-effective and rigid solutions that do not require a clean air supply.



➤ Figure 11: Hexapod System. (Image courtesy PI).

They have crossed-roller bearings that are generally the preferred choice due to their high stiffness and precision. A multitude of configurations is available, from compact voice-coil-driven systems providing 7mm XYZ-travel to larger stages covering 250mm or more. Direct drive motors, such as 3-phase motors, and voice coils are normally used, which help to deliver the highest performance and longest lifetime.

ACS Motion Control and drive modules

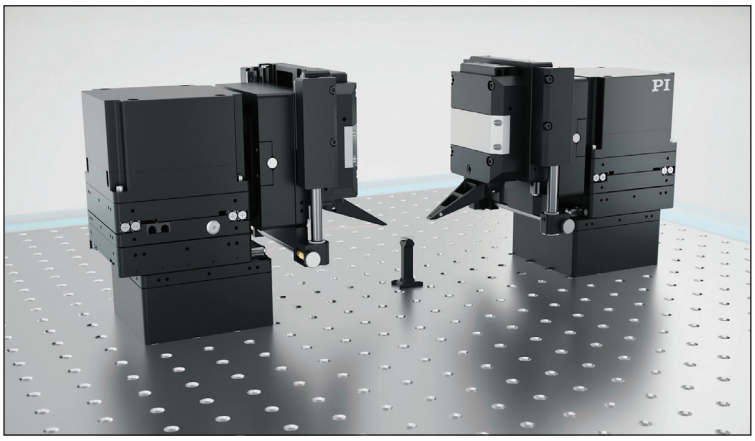
The photonics alignment systems pictured in Figures 11 and 12 operate with ACS Motion Controllers with integrated alignment firmware, as well as with system metrology that measures repeatability and accuracy with error mapping.

Additional alignment applications outside of SiPh

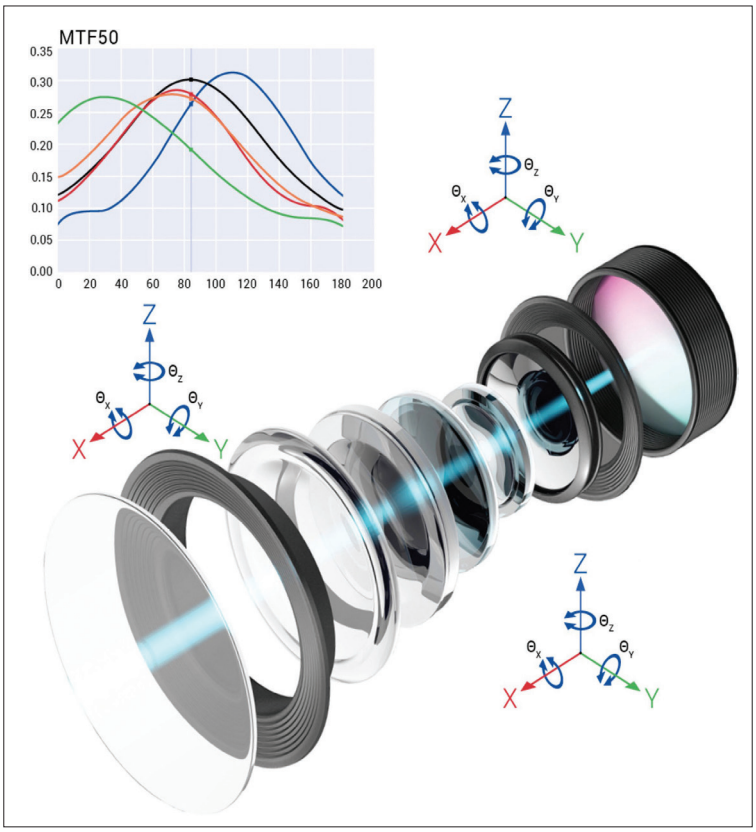
In addition to SiPh and fiber optics, precision alignment can increase yield and throughput in other applications with photonics-related components. In laser manufacturing, for example, bulk optics, cavities and diffractive elements need to be aligned to optimize the output beam's characteristics.This can be a time-consuming process, but automation can significantly reduce manufacturing time, and increase productivity and yield.

Similarly, high speed alignment can help manufacturers keep up with the ever-increasing demand for high resolution cameras which are, for example, used in smartphones and vehicle cameras. FMPA is a novel technology that drives this precision and speed, provides nanoscale

Modular alignment systems that consist of a combination of individual stages, known as serial kinematics, are guided by one of two designs: mechanical or air bearings. Air bearings are preferred as they provide precise, high dynamic motion that is free of friction, wear, and maintenance. Furthermore, air bearings do not generate particles as they neither contain lubricants nor are subject to abrasion caused by friction.



➤ Figure 12: PINovaLign42 Photonics Alignment System in a double-sided array alignment application. (Image courtesy PI).



autonomy to microrobots, and optimizes the position of multiple optical elements. With more sophisticated optical devices being produced each year, FMPA applications will continue to expand, ensuring precision positioning stays at pace with the rapidly evolving computing, imaging and photonics industries.

Summary

The process of aligning components is essential at every stage of photonic device production, starting with initial testing and grading on the wafer. This is traditionally a time-consuming and expensive task, and aligning hundreds or thousands of components on each wafer can be a daunting and, often, cumbersome task.

Traditional methods that take several minutes, such as aligning the vertical-cavity surface-emitting laser and photodiode channels on both ends of a photonic cable, may be valid for low volume, specialized items, but as demand exponentially increases, they become unfeasible. This is when automation becomes essential. Technologies such as FMPA clearly address this need, and are capable of meeting the current and future demands of the industry.

➤ Figure 13: Alignment of multiple lens elements is a complex task that can be accelerated by algorithms which are built into the controller firmware, along with novel mechanical alignment units, allowing the parallel alignment of multiple lenses in a single operation. Typically, the position of a lens system in front of a sensor is optimized using a calculation such as framewise MTF, which provides information about the sharpness of an image, with a better value signifying a better alignment. In other applications, such as single-photon quantum emitter studies, photon count is another signal that can be used to assess alignment. Basically, any sufficiently fast figure of merit that rises to a peak and then falls as a function of position can be optimized efficiently with FMPA. (Image courtesy PI).

ANGELTECH
CS INTERNATIONAL CONFERENCE PIC INTERNATIONAL CONFERENCE PE INTERNATIONAL CONFERENCE AP INTERNATIONAL CONFERENCE AIS ANGELTECH INNOVATE SUMMIT
SAVE THE DATE
12-14 APRIL 2027
Sheraton Brussels Airport Hotel
To find out more about our sponsor and speaker opportunities, contact us today on:
+44 (0) 2476 718970 or email:info@angeltech.net

semi™ | SEMICON® EUROPA

NOV 10-13, 2026
MESSE MUNICH, GERMANY

Europe's Next Era: Navigating the Perfect Storm

As global dynamics grow more complex, Europe faces both new, unprecedented challenges and unprecedented opportunities that will define its strategic position for decades to come.

INCREASE YOUR EXPOSURE!

Trade Show | Executive Programs | Networking | Conferences | Startup Showcase | Talent Initiatives

Contact: SEMICONEuropa@semi.org

TRANSFORM TOMORROW

Co-located with electronica

Turning water into a competitive advantage for the semiconductor industry



The semiconductor industry is entering a period of growth that would have looked implausible just a few years ago. The industry is on track to reach $1 trillion in annual revenue by 2030, and demand for AI-capable chips alone has grown by roughly 50% in the past three years.

BY DAVID HARRIS, GENERAL MANAGER, U.S., ULTRA-PURE WATER+, ECOLAB

THIS GROWTH presents a genuine opportunity for the industry -- more capacity, more capital investment, more advanced fabs coming online around the world. It is also placing new demands on a resource that rarely gets discussed with the same urgency as capital equipment or talent: water.

Ultrapure water (UPW) touches nearly every step of the chip manufacturing process, from wafer cleaning and wet etching to post-etch rinsing, chemical-mechanical planarization, and rinsing again. Get its quality even slightly wrong, and the consequence goes far beyond maintenance; it can cause production issues that can ripple through yield.

Historically, many fabs have managed that risk by treating ultrapure water largely as a linear input: draw it in, purify it, use it, discharge or treat what's left. That approach has served the industry well, but as fabs scale to meet AI-driven demand, it also leaves water reuse opportunity on the table. The fabs moving fastest today are the ones treating water not as a linear constraint to be managed, but as a circular resource that when optimized, can act as a lever for resilience, efficiency, and cost advantage.

Tighter tolerances, smarter measurement

Part of what's changing is straightforward physics. As process nodes shrink, the margin for contamination shrinks with them. A particle or trace metal ion that would have been irrelevant at older nodes can now interfere with pattern fidelity, gate integrity, or interconnect performance at the leading edge. That has pushed ultrapure water specifications for particles, dissolved metals, and organics to levels that are genuinely difficult to verify with conventional testing alone.

The encouraging part of this story is how quickly measurement technology is closing that gap. A growing number of fabs are moving from periodic grab sampling toward continuous, in-line sensing that tracks water quality parameter by parameter, in near-real time, at the point of use rather than only at the point of generation. That shift matters because it changes when a problem gets caught, not just whether it does.

–A deviation caught early is a process adjustment; caught late, after it has touched multiple tools and wafer lots, it becomes scrap and requalification. Continuous measurement gives engineering teams the lead time to correct course while it's still cheap to do so, turning what used to be a lagging risk indicator into a live operating input, tracked with the same discipline as temperature or gas purity.

A growing number of fabs are moving from periodic grab sampling toward continuous, in-line sensing that tracks water quality parameter by parameter, in near-real time, at the point of use rather than only at the point of generation.

Managing water as one connected system

Contamination risk doesn't originate at a single point in a fab's water system, which is part of why it has historically been hard to manage with a simple pass/fail checklist.

Ultrapure water quality depends on the interaction of feedwater characteristics, treatment performance, piping and storage materials, and the reclaim and reuse loops that more fabs are relying on to reduce freshwater withdrawal. A shift in incoming water conditions can propagate downstream in ways that don't show up until they appear in process data.

This is where opportunity exists. Ultrapure water, utility water, reclaim, and effluent are often managed as separate systems with separate owners inside a fab, even though they are hydraulically and chemically connected.

Fabs that instead treat the full water cycle as one integrated system, rather than a series of independent streams, gain visibility they didn't have before: a contamination source in one loop becomes something that can be traced and addressed before it quietly affects another.

That integration is also what allows a fab to pursue both higher purity and higher water reuse at the same time - two goals that can otherwise work against each other if they aren't engineered together from the start.

Water circularity as a scalable strategy, not a constraint

That integration is the foundation for a broader shift already underway across the industry: treating water circularity as a practical, scalable operating strategy rather than a sustainability add-on. The logic is simple and increasingly familiar to fab operators: reduce the water drawn into the process in the first place, reuse it across applications within the fab where quality permits, and recycle what's left through treatment that returns it to usable quality.

Handled this way, water circularity becomes a resilience strategy as much as an efficiency one. Fabs in water-stressed regions face real volatility: permitting timelines, allocation limits, and the reputational cost of being seen as a poor neighbor to a strained watershed. A fab that has already reduced its freshwater dependence and built reuse into its operations has more room to grow, more predictable operating costs, and a stronger case to make with regulators and communities alike. That is a competitive advantage, not just a compliance outcome, and it's increasingly a factor in where and how new capacity gets sited.

Growth is accelerating the timeline, not just the stakes

These technical and strategic questions are landing at a moment when the industry has little appetite for delay. Every fab coming online to meet AI-driven demand needs an ultrapure water system qualified and running reliably from day one. Once a system is qualified, operators have good reason to be reluctant to change it without a compelling case, which means decisions made at the design stage carry consequences for the life of the facility.

That puts a premium on getting the water strategy right early, rather than retrofitting it after commissioning. Fabs that bring water and process engineering into facility design conversations from the outset - rather than treating water as infrastructure to be solved once construction is underway - are the ones building in the flexibility to scale reuse, tighten purity control, and adapt to whatever the next process node demands, without having to re-engineer the system around a decision made years earlier.

Where the opportunity goes from here None of this has a single fix, and no single technology solves all of it.

What's changing is the recognition that ultrapure water quality, contamination control, and water circularity are no longer separate conversations; they're all part of the same system, and the fabs managing them together are pulling ahead. They treat water as one connected process rather than a set of independent streams. They invest in measurement that catches deviations before they reach the wafer. And they build reduce-reuse-recycle thinking into facility design from day one, rather than bolting it on later.

As the industry scales toward that trillion-dollar milestone, water will only become a more central variable in how fabs perform - not as a limiting factor to manage around, but as a lever the best-run fabs are already using to build resilience, improve efficiency, and create a durable competitive edge. The fabs that treat water that way now will be the ones best positioned to keep growing without missing a step.

David Harris is General Manager, U.S., Ultra-Pure Water+ at Ecolab. He has more than 40 years of experience in industrial water, reclamation, and wastewater treatment, including direct work in design, business management, and operations across the semiconductor, power, automotive, chemical, and refining industries.



Preventing downtime starts before a semiconductor fab starts operating



The semiconductor industry's race to add capacity will continue as AI and other advanced technologies increase chip demand. But adding capacity is only part of the equation. The goal is not simply to open a fab faster; it is to build one that can keep running uninterrupted.

BY TIMOTHY MCDONOUGH, SHERWIN-WILLIAMS CONSTRUCTION SOLUTIONS EXECUTIVE

HYPERSCALE AI data centers are rapidly expanding, driving extraordinary demand for the advanced semiconductor chips needed to train and deploy increasingly sophisticated models. As AI workloads expand, more servers are required, increasing demand for semiconductor chips. This places unprecedented pressure on the semiconductor industry to keep pace.

Available manufacturing capacity is struggling to keep up. According to a May 2026 report from the Center for a New American Security, capital expenditures (CapEx) for companies purchasing AI chips has grown exponentially faster than CapEx for companies producing AI chips. As of November 2025, demand for these chips was running at roughly three times the production output of the world's top manufacturers. This limit in chip supply chain is expected to remain through the end of 2026, at minimum.

With chip capacity already constrained by the increased demand, bringing new fabs online is only part of the challenge. Semiconductor facilities must maximize the availability of the capacity of their existing facilities. Unplanned maintenance, repairs or other disruptions can take production capacity offline in an environment where every hour of downtime matters, making reliability and continuous operation critical to both the individual fab performance and the industry's ability to close the broader supply gap.

Closing that gap will require new fabrication plants and the expansion of existing facilities. Major investments are already underway. Samsung and SK Hynix are investing more than $500 billion in new chip manufacturing facilities worldwide. Intel and TSMC have planned expansions of their US manufacturing capacity, investing more than $100 billion and $265 billion, respectively. But advanced fabs can take two to four years, and sometimes longer, to construct because of the scale, coordination and technical precision required.

The pressure is not only to build fabs faster. It is to build facilities that can operate continuously once production begins. Protective coatings are often an afterthought or an overlooked part of that equation. When planned early, the right coatings strategy can simplify construction, reduce rework and extend life cycle by protecting critical infrastructure against chemicals, corrosion and wear that can lead to premature maintenance and operation disruptions over time.



Right-size the coating systems

Semiconductor fabs use a wide range of acids, bases, solvents and specialty chemistries during silicon wafer production. Because the chemicals and exposure conditions vary across fabs and production areas, selecting chemical resistant coatings, or CRCs, can quickly become complicated.

That complexity often leads to one of two outcomes: an over-engineered system that adds unnecessary cost and construction steps or an under-engineered system that leaves steel, concrete and other substrates vulnerable to deterioration and premature failure.

More protection is not automatically better. The goal is to specify the right system for the actual demands of the environment. That means evaluating the chemicals in use, their concentrations and potential contact time, along with moisture, temperature, traffic, ultraviolet exposure and static-control requirements.

Those conditions can vary significantly across a fab, from secondary containment areas and chemical storage spaces to utility infrastructure, process-support

steel, floors and wastewater systems. Understanding what each asset will encounter allows teams to match the level of protection to the actual exposure rather than applying a one-size-fits-all approach across the facility.

Defining those conditions early allows project teams to develop application-specific systems instead of relying on broad assumptions. A clear and repeatable selection process can simplify specifications across the fab, reduce late-stage changes and create a scalable approach for future expansions. Just as importantly, it can help reduce premature coating failures that may require repairs or maintenance around active production areas later.

Move work off-site

Another way protective coatings can support schedule certainty and reduce construction risk is to shift appropriate scopes away from the jobsite. With shop application, coatings are applied to steel, pre-cast concrete or other components in a controlled environment before they arrive on-site. This approach reduces weather-related delays and provides greater control over surface preparation, application conditions and quality assurance. It also requires fewer crew members and less equipment on-site, helping improve safety and reducing congestion among specialized trades.

These shopcoated components can arrive ready for installation and touch-up, allowing fabrication, coating and site preparation to progress at the same time. Creating these parallel workstreams simplifies sequencing and limits the bottlenecks that can slow a megaproject.

The coating system must be specified to withstand loading, transportation, erection and weather exposure. When those requirements are addressed correctly, shop-applied coatings can support faster construction while delivering durable asset protection. This doesn't just apply at initial construction, it also is relevant for future expansions. reducing on-site coating activities can also help owners add capacity without limiting work around active production environments.

Plan for uptime early

Coating decisions are often treated as a finishing detail and pushed later in the project. By then, the construction sequence, procurement plan and access requirements may have already been established. Any mismatch between the coating specification and the broader build plan can create delays, add change orders and rework during construction, or leave owners with systems that are more difficult to maintain once the fab is operating.



Early collaboration helps prevent those problems. A knowledgeable coatings partner can work with owners, architects, engineers, general contractors and fabricators to align system selection with fabrication, delivery, installation, cure windows and inspection requirements, while also considering how those systems will be maintained, repaired or expanded once the facility is operational.

That partner should provide more than products. Support may include specification development, application guidance, contractor training, field services and more. The value is in identifying potential roadblocks before they affect the schedule as well as helping teams make faster, more informed decisions if and when conditions change.

Establishing a repeatable coating strategy during the initial build also provides a stronger foundation for maintenance, expansions and fast return-to-service work later. Owners and maintenance teams have a clearer understanding of which systems protect specific assets, the conditions those systems were designed to withstand and how they should be maintained or repaired over time. When an expansion or maintenance event occurs, teams are not starting the coating-selection process from scratch, simplifying repairs and returning affected areas to service faster.

Build fast, operate continuously

The semiconductor industry's race to add capacity will continue as AI and other advanced technologies increase chip demand. But adding capacity is only part of the equation. The goal is not simply to open a fab faster; it is to build one that can keep running uninterrupted.

Once production begins, coating performance becomes an uptime consideration. Chemical exposure, moisture, corrosion and everyday wear can degrade steel, concrete and other critical infrastructure over time. When deterioration requires repair or maintenance in an operating fab, the work can create disruption around active production lines, which can cause quality issues or other financial impacts. Properly specified coating systems can help reduce that risk by protecting assets against the conditions that cause premature deterioration and minimize the likelihood of unplanned repair or rework. And when maintenance is eventually required, a well-planned coating system can support faster return to service.

Right-sized CRC systems, shop-applied coatings and early collaboration can turn coatings from a late-stage specification into a tool for schedule certainty, asset reliability and long-term operational performance. In semiconductor manufacturing, protecting uptime begins long before the first wafer is processed.

From water conservation to climate action:

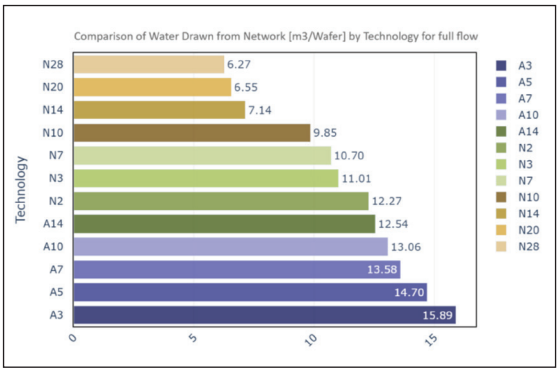
SCREEN's holistic path to sustainable semiconductor manufacturing

The semiconductor industry faces a defining challenge. As demand for advanced chips accelerates, manufacturers must simultaneously increase production capacity while dramatically reducing their environmental footprint. Water consumption, wastewater generation, resource efficiency, and carbon reduction have become strategic priorities across the semiconductor value chain.

BY LAURA AMATI, SCREEN

FOR SCREEN, headquartered in Kyoto, Japan, water stewardship is more than a technical challenge. The company's sustainability philosophy is deeply influenced by Kyoto's cultural heritage and long-standing respect for water as a vital natural resource. This perspective has helped shape SCREEN's commitment to developing technologies that reduce environmental impact while maintaining the highest levels of process performance.

At SCREEN Semiconductor Solutions, sustainability is no longer treated as a standalone environmental initiative. It has become a core element of product development, manufacturing operations, and industry collaboration. By combining equipment-level innovation, digitalization, water management expertise, and ecosystem partnerships, SCREEN is helping semiconductor manufacturers move toward a more sustainable future.



➤ Figure 1. Water Withdrawal from Network (m^3/wafer) per technology node.

Water: One of the Semiconductor industry's most critical resources

Modern semiconductor manufacturing is highly water intensive. A leading-edge semiconductor fab can consume tens of millions of liters of ultrapure water (UPW) per day. [1] As device architectures become increasingly complex, the number of wet process steps continues to grow, creating additional pressure on water resources as shown in Figure 1.



➤ Figure 2. Overview of SCREEN Hikone plant.

Advancing sustainable water management through tool-facility integration

Historically, water conservation efforts in semiconductor manufacturing focused primarily on centralized wastewater treatment and reclamation systems. While effective, these end-of-pipe approaches alone are no longer sufficient to meet increasingly ambitious sustainability objectives.

SCREEN promotes a more integrated strategy that combines tool level optimization with facility-wide water management. This approach aligns closely with the principles outlined in SEMI's F98 guideline for water reuse in semiconductor manufacturing, which promotes centralized, decentralized, and hybrid water reuse strategies. [2]

To support this evolution, SCREEN has partnered with FTD Solutions to develop advanced digital twin capabilities for semiconductor fabs. Through the

Water Management Application (WMA®), facilities engineers and water management staff can create detailed digital representations of water systems, enabling:

- Real-time monitoring of water and chemical flows
- Data-driven optimization of utility systems
- Identification of conservation opportunities
- Sustainability performance tracking and reporting
- Scenario modeling for future water-reduction initiatives

SCREEN has already implemented WMA® modeling at its Hikone R&D and manufacturing site in Japan. The project provided a comprehensive site-wide understanding of water flows, chemical balances, infrastructure connections, and operational performance. These insights are supporting the development of future water-conservation strategies and more sustainable fab designs.

By integrating equipment data directly into facility-level digital twins, SCREEN aims to demonstrate how individual tool configurations can influence overall fab sustainability performance.

Reducing water consumption at the source

In addition to facility optimization, SCREEN continues to develop innovations within its wet processing platforms to reduce water use directly at the source.

One example is the implementation of advanced recirculation technologies for hot ultrapure water (HUPW) and ozonated ultrapure water systems. Instead of discharging valuable water during periods of reduced tool utilization, recirculation allows water quality to be maintained while minimizing wastewater.

Another example is SCREEN's enhancement of the eFLOW® system used on its cleaning platforms. By reducing unnecessary drain flow during operation, the solution significantly lowers DI water consumption while maintaining process stability. Such improvements illustrate how relatively small design changes can generate meaningful environmental benefits when deployed across large-scale manufacturing operations.

These initiatives support SCREEN's broader objective of helping customers reduce freshwater requirements without compromising yield, throughput, or reliability.

Transforming wastewater into a resource

Sustainability requires more than conservation. It also requires maximizing the value of resources already entering the manufacturing process. A growing area of focus is the treatment and recovery of wastewater streams generated by advanced semiconductor manufacturing, particularly those containing isopropyl alcohol (IPA). Rather than managing these streams exclusively through centralized treatment systems, SCREEN is exploring segregation-based solutions that separate wastewater at the tool level and enable more efficient recovery.



➤ Figure 3. Single Wafer Cleaner SU-3200.

Using membrane-based separation technologies, IPA-rich wastewater can be concentrated and recovered, opening opportunities for reuse while simultaneously reducing wastewater volumes. This shift from wastewater treatment to resource recovery supports the industry's transition toward more circular manufacturing models and helps reduce both environmental impact and operational costs.

Collaboration across the semiconductor ecosystem

SCREEN recognizes that no single company can solve the semiconductor industry's sustainability challenges alone. Meaningful progress requires collaboration across equipment suppliers, device manufacturers, research organizations, and industry associations.

The future of sustainable semiconductor production will not be achieved through a single breakthrough. It will be built through thousands of improvements across equipment, facilities, and industry partnerships. By integrating sustainability into every level of its operations and innovations, SCREEN Semiconductor Solutions is helping shape that future.



Figure 4. SCREEN corporate purpose.

Semiconductor Climate Consortium
FOUNDING MEMBER

To support industry-wide climate action, SCREEN became a founding member of the Semiconductor Climate Consortium (SCC), an initiative established by SEMI. The consortium's mission is to accelerate decarbonization across the semiconductor value chain through collaboration, transparency, and ambitious collective action.

As stated by the consortium:
"The mission of the Semiconductor Climate Consortium is to speed the decarbonization of the semiconductor value chain through collaboration, transparency and ambition. By leveraging accumulated knowledge and innovative technology and pooling our resources, we identify the most pressing problems and work together to address them on an industry scale."

Through its participation, SCREEN contributes to the development of practical solutions that address climate-related challenges facing semiconductor manufacturing while promoting alignment across the industry. The company continues to support initiatives that reduce greenhouse gas emissions, improve resource efficiency, and encourage transparent sustainability reporting throughout the semiconductor ecosystem.

SCREEN also participates in imec's Sustainable Semiconductor Technologies and Systems (SSTS) research program. Launched in 2021, SSTS brings together major semiconductor industry stakeholders to reduce the environmental impact of chip manufacturing through collaborative research and lifecycle assessment methodologies.

Leveraging imec's net-zero lifecycle assessment platform, the program evaluates environmental impacts across semiconductor manufacturing processes and identifies pathways toward more sustainable production. [3]

Participation in SSTS provides SCREEN with opportunities to collaborate on next-generation sustainability solutions while helping establish industry best practices for environmental performance measurement and improvement.

Looking ahead

As semiconductor manufacturing continues to expand globally, sustainability will increasingly become a competitive differentiator and operational necessity. Water scarcity, climate targets, energy use, and circular resource management will require innovative thinking across every stage of the manufacturing process.

SCREEN's strategy combines three critical elements:

- **Tool-level innovation** to reduce water and resource consumption.
- **Digital transformation** through facility-wide modeling and optimisation.
- **Industry collaboration** through organisations such as the Semiconductor Climate Consortium and imec's SSTS program.

Together, these efforts reflect SCREEN's vision of creating technologies that enable both advanced semiconductor manufacturing and environmental responsibility.

The future of sustainable semiconductor production will not be achieved through a single breakthrough. It will be built through thousands of improvements across equipment, facilities, and industry partnerships.

By integrating sustainability into every level of its operations and innovations, SCREEN Semiconductor Solutions is helping shape that future.

FURTHER READING

- [1] World Economic Forum, Semiconductor manufacturing and big tech's water challenge | World Economic Forum, Jul 19, 2024.
- [2] SEMI F98-0923, Guide for Water Reuse in Semiconductor Industry, SEMI, 2023.
- [3] imec NetZero, imec.netzero, September, 2026

To promote your products and services contact: **Shehzad Munshi**
T: +44 (0)1923 690 215 **E:** shehzad.munshi@angelbc.com

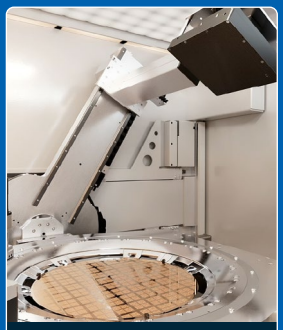
CLEAN AIR

CAMCARB XG
Protecting your facility from airborne molecular contamination.
- High filtration performance
- Reduced energy costs
- Longer operational lifetime



camfil
www.camfil.com/camcarb-xg

INSPECTION SYSTEMS



Powerful inspection solutions for wafer and chip level packaging.

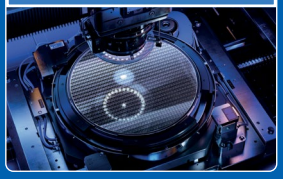
comet yxlon.comet.tech

COMPONENTS

Greene Tweed

WHEN IT CAN'T FAIL, CHOOSE GREENE TWEED.
- SEALS
- CONNECTORS
- COMPONENTS

VISIT **GTWEED.COM** TODAY



ETCH

Replace Your Old Legacy Ashing Systems with ENVIRO™



- Replaces up to 6 legacy tools
- >12µm/min ashing rate
- 245 WPH
- Lowest CoO ashing system

Contact ULVAC, and see how Enviro ashing tools will add horsepower to your 200 mm FAB.

ULVAC
ULVAC GmbH
+49 – 89 – 96 09 09 0
ulvac@ulvac.de | www.ulvac.eu

FLUID SYSTEMS



Performance Today, Innovation Tomorrow

Find the fluid system solutions you need:
www.swagelok.com/industries/semiconductor

Swagelok

GAS DELIVERY SYSTEMS

APPLIED ENERGY SYSTEMS

Your world of discovery, **powered by our delivery.**



Ultra high purity gas delivery solutions for rigorous purity and precision needs.

www.appliedenergysystems.com

GAS PURIFIERS

ARM PURIFICATION
BY APPLIED ENERGY SYSTEMS

SO MUCH MORE THAN A COMPONENT.



Advanced point-of-use, micro-bulk and bulk **gas purifier solutions**

www.appliedenergysystems.com/pure

HYDROGEN GENERATORS

nel

Nel Hydrogen is the recognized industry leader of Alkaline and PEM water electrolysis. We provide solutions that meet global hydrogen requirements for semiconductor and thin film manufacturing worldwide.



+1.203.949.8697
info@nelhydrogen.com
www.nelhydrogen.com

GAS ANALYSERS



Where ultra-high-purity gas analysis meets semiconductor precision

Discover Gen-7 DF-700 series
servomex.com/gen7

SERVOMEX
a spectris company

PLASMA TECHNOLOGY

Trymax®

When Plasma Matters



www.trymax-semiconductor.com

TEST

SPEA
YOUR BEST WAY TO TEST

DOT Series
Next-generation Semiconductor Testers



Discover how to save **up to 90%** on your cost of test

www.spea.com/savewithDOT

VACUUM EQUIPMENT



Powerful Turbopumps HiPace 2300 with Pumping Speeds of up to 2050 l/s

PFEIFFER VACUUM
Your Success. Our Passion.

T 800-248-8254
www.pfeiffer-vacuum.com



YOUR SUBFAB.

How much is it taking from you?

The true cost of ownership isn't only measured in kilowatts.

It's found in maintenance hours, unplanned downtime, service visits and the time your team spends keeping performance on track.

Ganymede helps reduce the total cost of ownership of your vacuum infrastructure through lower energy consumption, optimised maintenance planning and data-enabled performance management.

So while your operation spends less, your team spends less energy.

Lower costs, more control.

Discover how Ganymede is changing the economics of vacuum.

EDWARDS