



Chip Scale Review®

ChipScaleReview.com

The Future of Semiconductor Packaging

Volume 30, Number 3

Summer 2026

Hybrid bonding: A critical enabler of next-gen 3D device architectures

- Through glass via (TGV) etch comes to maturity for glass core applications
- Bumpless flip-chip (BFC) as a next-generation interconnect technology for the AI era
- Enabling high-performance 3D integration through advanced microbump technology
- Potential applications of Cu-Cu hybrid bonding to heterogeneous integration of chiplets
- Achieving faster ramps and lower costs with yield prediction technology in FOPLP lithography

CONTENTS

Volume 30, Number 3
Summer 2026



*In hybrid bonding, tight alignment and overlay accuracy between the wafers is required to achieve good electrical contact between the interconnected devices on the bonded wafers. Constant reduction in pitches is fueling tighter bonding specifications. Shown here is the SmartView® NT3 Aligner, which achieves sub-50nm alignment accuracy on 300 mm wafers, supporting advanced 3D applications such as memory and logic stacking, AI accelerators, and heterogeneous integration.
Cover image courtesy of EV Group*

FEATURE ARTICLES

- 5 Enabling high-performance 3D integration through advanced microbump technology**
By Calvin Lee [[ASE US](#)]
- 11 Bumpless flip-chip (BFC) as a next-generation interconnect technology for the AI era**
By Rabindra N. Das, Dan Oates, Alexander Wynn, David Kim, Steven Weber, Mollie Schwartz
[[Advanced Technology Division MIT Lincoln Laboratory](#)]
- 18 Hybrid bonding: A critical enabler of next-gen 3D device architectures**
By Tobias Wernicke, Bernd Dielacher [[EV Group](#)]

CONTENTS

Chip Scale Review®

ChipScaleReview.com

The Future of Semiconductor Packaging

STAFF

Kim Newman

Publisher

knewman@chipscalereview.com

Lawrence Michaels

Managing Director

Editor-in-Chief

lmichaels@chipscalereview.com

Debra Vogler

Senior Technical Editor

debravogler@me.com

SUBSCRIPTION—INQUIRIES

Chip Scale Review

All subscription changes, additions, deletions to any and all subscriptions should be made by email only to subs@chipscalereview.com

Advertising Production Inquiries:

Lawrence Michaels

lmichaels@chipscalereview.com

Copyright © 2026 Haley Publishing Inc.

Chip Scale Review is a registered trademark of Haley Publishing Inc.
All rights reserved.

Subscriptions in the U.S. are available without charge to qualified individuals in the electronics industry.
Chip Scale Review is published four times a year with issues in Winter, Spring, Summer and Fall.

P.O. Box 2165
Morgan Hill, CA 95038
E-Mail: info@chipscalereview.com

Printed in the United States



FEATURE ARTICLES *(continued)*

24 Achieving faster ramps/lower costs with yield prediction technology in FOPLP lithography

By John Chang, Jian Lu, Timothy Chang [*Onto Innovation*]

32 Potential applications of Cu-Cu hybrid bonding to heterogeneous integration of chiplets

By John H. Lau [*Unimicron Technology Corporation*]

40 Through-glass via (TGV) etch comes to maturity for glass-core applications

By Venugopal Govindarajulu, Zia Karim [*Yield Engineering Systems (YES)*]

DEPARTMENTS

INDUSTRY EVENTS

47 IMAPS Symposium 2026: The premier event for microelectronics advanced packaging

By Karthik Dhandipandi [*Cerebras Systems; General Chair IMAPS 59th International Symposium on Microelectronics*]

49 35th SWTest Conference and EXPO: Sold out and new record attendance!

By Jerry Broz, PhD [*General Chair, SWTest US and SWTest Asia Conferences, and VP at Delphon Industries*]

53 2026 IEEE ECTC: Third consecutive record attendance shows advanced packaging is the enabler for HPC, AI and more

By Michael Mayer [*Associate Professor, University of Waterloo; General Chair 76th ECTC*]

Hybrid bonding: A critical enabler of next-gen 3D device architectures

By Tobias Wernicke, Bernd Dielacher [EV Group]

The rapid advancement of microelectronics is placing increasing pressure on conventional device architectures to deliver higher performance while maintaining efficiency and scalability. Traditional monolithic scaling is no longer sufficient. As a result, the industry is moving towards new integration schemes combining individually optimized components via advanced 3D stacking and heterogeneous integration approaches.

The technological shift towards new integration schemes is driven in large part by the rapid rise of artificial intelligence (AI), cloud computing, and other data-intensive applications. These workloads place unprecedented demands on compute performance, memory bandwidth, and energy efficiency. Modern data centers are scaling aggressively, and applications such as large language models and real-time analytics require tightly-coupled compute and memory architectures. At the same time, the need to balance performance with power consumption and thermal constraints is becoming more critical than ever. Consequently, system architectures must evolve to support higher data movement efficiency and reduced latency, enabling sustained performance improvements in an era defined by exponential data growth.

In this context, hybrid bonding is emerging as a key technology to meet these escalating system requirements. It provides ultra-fine-pitch connectivity well beyond the capabilities of traditional micro-bump solutions and has become a key enabler of advanced 3D and heterogeneous integration. But what is hybrid bonding? Unlike conventional bonding techniques, hybrid bonding uniquely combines dielectric-to-dielectric fusion bonding with metal-to-metal thermocompression bonding (TCB), usually copper to copper, within a single process step.

Plasma-activated fusion bonding allows the joining of two dielectrics at room temperature through surface-to-surface interactions, primarily driven by capillary forces and secondarily by Van der Waals forces. The initial state, called “pre-bond,” is strong enough to allow wafer handling but requires a subsequent annealing step to form a permanent bond. During this temperature treatment, the pre-bond is transformed into a covalent bond via a hydrolysis reaction between surface oxides and hydroxyl (OH) groups. This bonding mechanism is well described by the Gösele model [1]. To reduce the annealing temperature to a complementary metal-oxide-semiconductor (CMOS)-compatible range (<400°C), the substrates are plasma activated before the pre-bonding. This activation modifies surface hydrophilicity, which ultimately increases the density of reactants for the hydrolysis reactions, most notably hydroxyl (OH) groups [2].

The electrical connection during a hybrid bond is established via a very local metal-to-metal TCB. The driving mechanism during TCB is solid-state interdiffusion. Because there is no liquid metal phase involved, no related wafer-to-wafer alignment shift occurs.

The process of record to create hybrid-bondable structures is the damascene process (Figure 1). While the patterning process can vary significantly across different damascene integration schemes (e.g., dual-damascene, single-damascene), the preparation of the bonding interface

surface always involves an over-polishing step, resulting in a recessed metal surface relative to the bonding dielectric. Because both parts of the hybrid bonding—the plasma-activated fusion bonding and the TCB—require a temperature step, this over-polish is required to compensate for the mismatch in thermal expansion between the metal and the dielectric materials. In addition, to ensure reliable mechanical contact between both materials, the degree of metal recess is a critical parameter. Typically ranging from 2-4nm, the metal recess must be controlled across the entire substrate with sub-nanometer precision.

When looking at roadmaps (Figure 2), particularly in terms of pitch scaling and associated wafer-to-wafer overlay requirements, the primary challenge in current hybrid bonding processes becomes very obvious: How can wafer-to-wafer overlay be controlled and further improved? Considering the latest reference data [5], the success of a hybrid bond in terms of wafer-to-wafer overlay is typically evaluated based on the maximum misalignment vector measured across the wafer. However, to optimize wafer overlay performance, it is essential to understand the individual contributors to misalignment and their respective origins.

The EVG overlay model analyzes an overlay measurement, which represents the vector-based misalignment between top-to-bottom alignment marks at multiple locations across the wafer

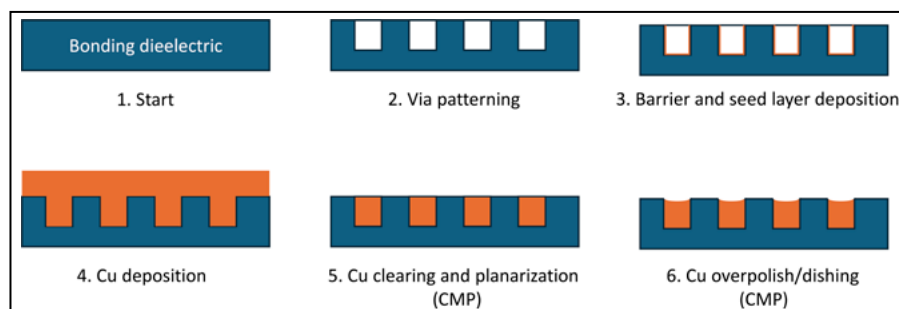


Figure 1: Simplified damascene process flow. SOURCE: EVG

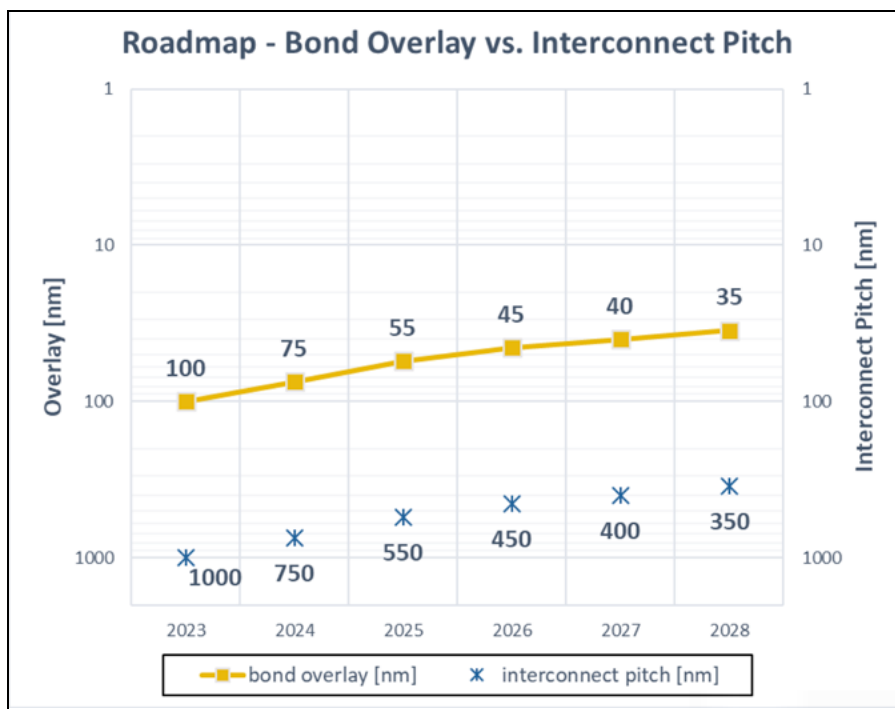


Figure 2: Roadmap bond overlay vs. interconnect pitch [3]. SOURCE: EVG

(“Overlay”), and separates them into overall linear (“Alignment”) and nonlinear (“Distortion”) components (Figure 3).

The alignment component describes the global wafer-to-wafer shift and comprises four sub-components: Translation in the x-direction, translation in the y-direction, rotation, and wafer-to-wafer scaling. By subtracting these linear components from the measured overlay, the remaining residuals—referred to as distortion—can be isolated and analyzed.

Origin of alignment errors

Global shifts between wafers in x- and y-directions are straightforward to understand and can be directly corrected using standard wafer-to-wafer optical alignment systems. The scaling component, however, can originate from various sources, such as environmental differences during top- and bottom-production, respectively, and different stress-induced wafer shapes.

It is important to mention that the bonding process itself also introduces a scaling component. This is primarily attributed to the point-contact initiation of the bonding process via a bond pin, which triggers the fusion bond process (Figure 4). During the bonding process, the bond pin locally pushes the center of the top wafer towards the bottom

wafer while the outer regions are still being held by vacuum. This results in a radial stretching of the top wafer, which is transferred to the bottom wafer as the bond wave propagates. Various techniques are available in

industry to compensate for this inherent scaling effect, as illustrated in Figure 5. Because the scaling induced by the bonding process exhibits a strong linear dependence on the curvature of the top wafer, a highly effective compensation method is to intentionally induce a matching curvature on the counter-wafer, thereby canceling the inherent scaling effect.

Origin of distortion errors

When analyzing distortion behavior and its origins in hybrid bonding, a distinction must be made between blank and patterned wafers. As shown in Figure 6, the bond wave propagation of a patterned wafer differs significantly compared to a blank wafer, adopting a more square-shaped profile. This behavior arises because the bond wave of the dielectric regions must propagate around the embedded metal features, which are typically arranged in orthogonal (horizontal-vertical) layouts. As a result, the bond wave front is constrained by the metal pattern geometry, which translates into different bond wave shapes. Ultimately, this behavior has a direct impact on the overlay.

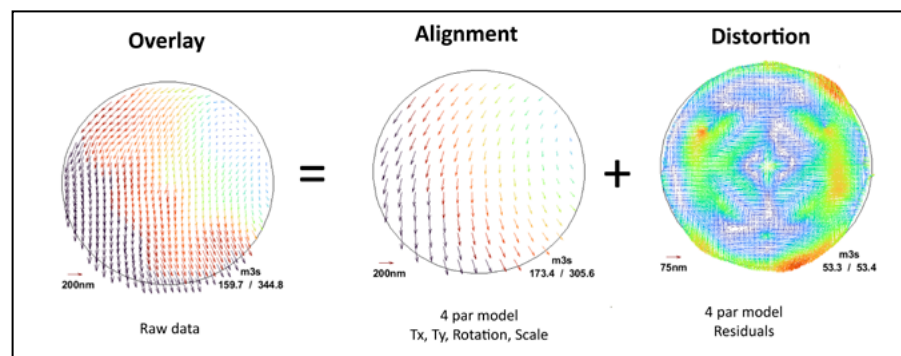


Figure 3: Example of overlay breakdown. SOURCE: EVG

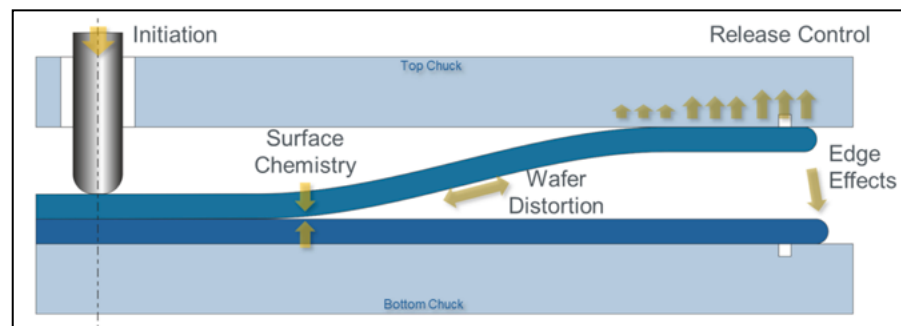


Figure 4: Scaling and distortion dynamics during the fusion bonding process. SOURCE: EVG

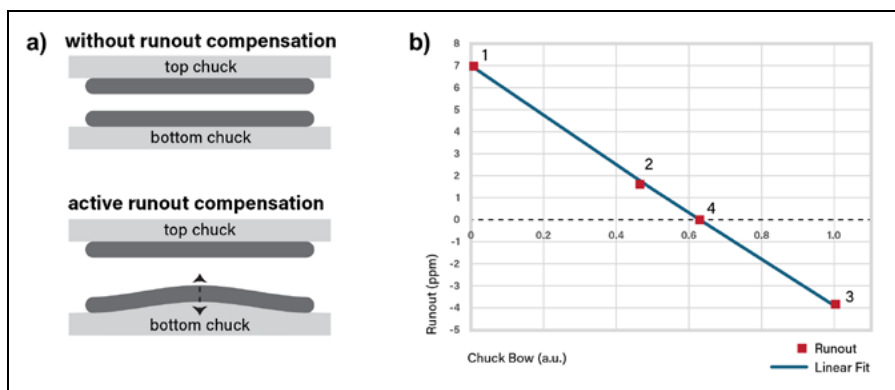


Figure 5: a) Schematic illustration of scaling (runout) compensation in wafer bonding; and b) Linear relationship between chuck bow and runout [4]. SOURCE: EVG

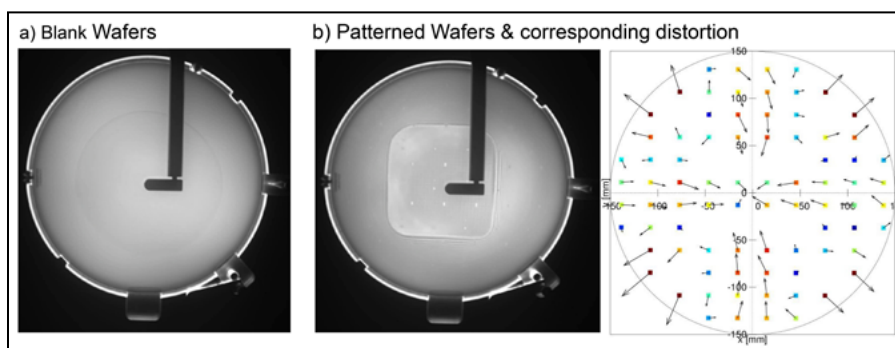


Figure 6: Bond wave propagation behavior of: a) Blank vs. b) Patterned wafers. SOURCE: EVG

Benefits of hybrid bonding

Hybrid bonding offers a scalable pathway to extend system-level performance beyond the limits of traditional device scaling. It enables each functional chip to be independently optimized, leveraging the most suitable process node or material while maintaining tight system-level coupling. As a result, designers can move beyond monolithic scaling and adopt more flexible, performance-driven integration strategies. These approaches are already being deployed in applications such as AI accelerators and high-performance computing (HPC), where memory bandwidth and latency represent critical bottlenecks. In this context, hybrid bonding enables the close integration of logic and memory, supporting advanced architectures such as high-bandwidth memory (HBM) stacks and emerging logic-on-memory concepts.

The developments noted above align closely with imec's CMOS 2.0 vision, which extends 3D integration towards system-level partitioning into distinct tiers, reconnected through dense vertical interconnects. These architectures require extremely fine pitches for logic-to-logic and memory-to-logic stacking. Recent

results from imec and EV Group highlight the state of the art in hybrid bonding [5], demonstrating wafer-to-wafer hybrid bonding at a 200nm interconnect pitch with sub-40nm alignment accuracy across full 300mm wafers, representing a record level of precision (Figure 7). This breakthrough, enabled by optimized materials, chemical-mechanical polishing (CMP) processing, and advanced bonding equipment, marks a key milestone toward the interconnect densities required for CMOS 2.0 and future 3D system architectures. Looking further ahead, hybrid bonding is also expected to support future device architectures such

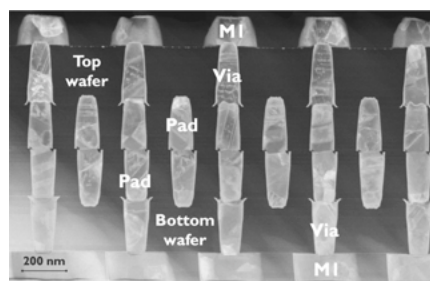


Figure 7: Transverse electromagnetic mode (TEM) image of daisy chain structures on a 200nm hexagonal pad grid with equal hybrid pad size and 25% designed Cu density [5]. SOURCES: imec/EVG

as complementary field-effect transistor (CFET), where vertical stacking at the transistor level will impose even more stringent requirements on nanoscale interconnect definition and alignment.

New hybrid bonding challenges

The challenges of these scaling and integration strategies extend beyond wafer-to-wafer overlay requirements and include material-related issues, such as thermal and electrical properties, discussed below.

In terms of thermal properties, particularly in the context of logic-memory integration, increasingly stringent temperature budget must be maintained to avoid issues such as unwanted diffusion or excessive thermomechanical stress. This temperature budget can be below 200°C, which is significantly below the commonly used annealing temperatures for hybrid bonding (around 300°C). One promising approach to achieve bonding at reduced temperatures is copper grain-engineering, specifically, fine-grain copper, often referred to in literature as nano-crystalline copper [6]. The modification of the grain structure is typically achieved during the copper electroplating by introducing different additives during deposition that impact the grain-growth behavior. Because the metal bonding mechanism in hybrid bonding is based on diffusion, the use of fine-grain copper is motivated by the increased density of grain boundaries, which act as fast diffusion pathways. Diffusion along grain boundaries occurs at significantly higher rates than bulk (volume) diffusion. This allows interconnect formation similar to “conventional copper,” but at significantly reduced temperatures. During annealing, fine-grain copper recrystallizes resulting in low electrical resistivity.

Regarding the electrical properties of hybrid bond interconnects, significant research efforts are focused on sub-10nm via dimensions, as required for advanced logic applications, such as the hybrid-bonded CFET architecture. While copper-based hybrid bonding is well established, scaling copper to these dimensions becomes increasingly challenging for two main reasons. First, copper can easily diffuse into silicon, necessitating the use of a robust diffusion barrier within the interconnect structure. As dimensions shrink, this barrier must scale accordingly. Second, the electrical resistivity of thin-film copper increases exponentially for

film thicknesses approaching the mean free path of copper (~40nm), primarily due to enhanced surface and grain boundary electron scattering. Therefore, alternative materials are being investigated—the most promising being ruthenium—particularly when grown epitaxially, as it is well suited for front-end integration.

As shown in **Figure 8**, ruthenium, especially epitaxial ruthenium, exhibits favorable resistivity for thin films below 10nm. In addition, ruthenium interconnects do not require diffusion barriers, like for copper interconnects. This enables a significantly higher fraction of the interconnect cross section to be occupied by the conducting material, rather than by auxiliary layers such as barriers [7]. Consequently, ruthenium is considered one of the most promising candidates for advanced interconnect scaling in hybrid bonding applications.

An equally important aspect of

advanced hybrid bonding integration schemes is the need to enable the extreme wafer thinning of these architectures. As interconnect pitch scales into the sub-micron regime, the entire stack must become thinner, flatter, and more precisely controlled. This shifts temporary bonding from a supporting process step to a critical technology driver on its own. An enabler in this domain is EVG's LayerRelease™ technology, an infrared (IR) laser debonding technology based on silicon carriers and inorganic release layers that allows fusion bonding to the carrier—providing superior rigidity during device thinning. In addition, its inorganic nature is front-end compatible and supports high-temperature processing steps directly on the carrier. IR laser debonding has emerged as a particularly effective solution for next-generation integration schemes that require extreme thinning in advanced process flows.

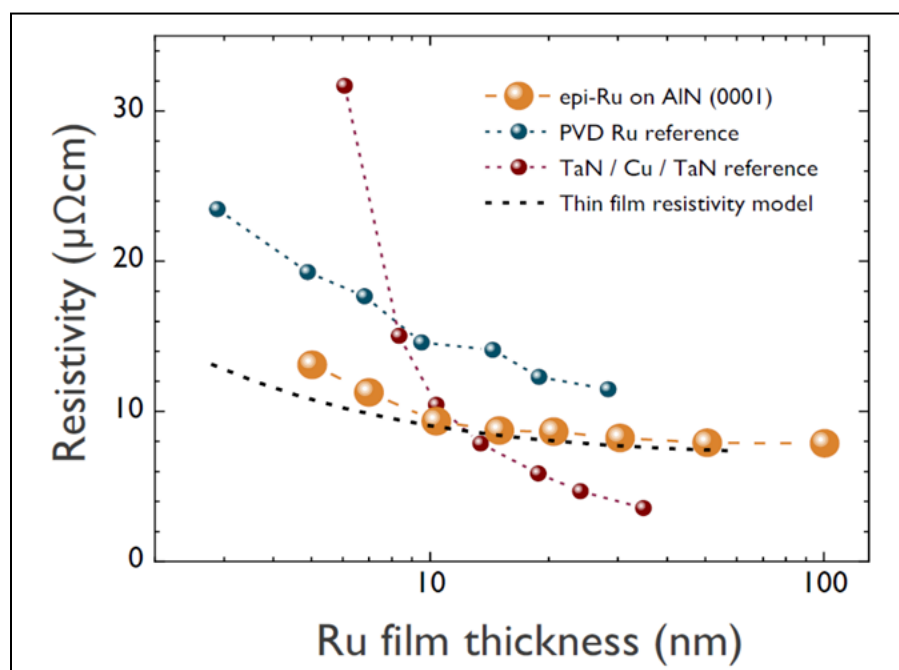


Figure 8: Resistivity of ruthenium and other metals related to film thickness [7]. SOURCE: Used with permission of IEEE

Summary

As the semiconductor industry continues to push beyond the limits of conventional scaling, hybrid bonding is rapidly evolving into the dominant interconnect technology for advanced integration. While wafer-to-wafer overlay is constantly optimized, new materials and processes are being introduced to enable new technology nodes and integration schemes.

References

1. Q.-Y. Tong, U. Gösele, *Semiconductor Wafer Bonding Science and Technology*, Wiley, (1998).
2. N. Rauch, et al., "A model for spectroscopic ellipsometry analysis of plasma activated Si surfaces for direct wafer bonding," *Applied Phys. Lett.* 121(8), 081603 (2022).
3. T. Wernicke, et al., "Wafer bonding as next generation scaling booster," *International Electron Devices Meeting (IEDM)*, San Francisco, CA, USA (2023).
4. T. Plach, et al., "High accuracy aligned wafer bonding for wafer-level integration," *ECS Trans.*, 86 (5) 145-158 (2018); DOI 10.1149/08605.0145ecst.
5. Press Release: "Imec and EV Group demonstrate wafer-to-wafer hybrid bonding with 200nm interconnect pitch and record high overlay accuracy," May 28 (2026).
6. M. Maubert, et al., "Copper microstructure optimization for fine pitch low temperature Cu/SiO₂ hybrid bonding," *IEEE 74th Electronic Components and Technology Conf. (ECTC)*, Denver, CO, USA (2024).
7. C. Adelman, et al., "Layer transfer of epitaxial Ru by metal-metal bonding: Towards single-crystal interconnects," *IEEE 76th ECTC*, Orlando, FL, USA (2026).



Biographies

Tobias Wernicke is Process Technology Manager for wafer bonding at EV Group, St. Florian, Austria, where he oversees all wafer bonding process developments. Tobias holds a Master's in Material Science and Nanotechnology at the Friedrich-Alexander University of Erlangen. Email T.Wernicke@evgroup.com

Bernd Dielacher is Director of Business Development at EV Group, St. Florian, Austria, where he leads global business development efforts, contributes to technology roadmaps, and supports new product developments for EVG's core technologies. Bernd holds a Master's in Microelectronics from Vienna University of Technology, and a PhD in Biomedical Engineering from ETH Zurich.