Advanced Packaging and the New Scaling Walls

September 17, 2026 | Chetan Arvind Patil, Marvell Technology

Estimated reading time: 4 minutes


Chetan Patil

Advanced packaging is becoming the foundation for semiconductor scaling in the AI era. As monolithic processors encounter practical limits around reticle size, yield, memory bandwidth, power, and cost, the industry is moving toward architectures built from multiple compute dies, HBM stacks, I/O chiplets, and heterogeneous packaging solutions.

The result is that semiconductor scaling is no longer determined only by how much functionality can be integrated into a single die, but by how much silicon can be integrated efficiently within a package.

This transition, however, does not eliminate the limits that constrain traditional semiconductor scaling. It simply relocates them. As packages become larger and more heterogeneous, the industry faces a new collection of interrelated constraints involving package dimensions, assembly yield, test coverage, thermal density, power delivery, mechanical stability, and manufacturing economics.

Advanced packaging and heterogeneous interconnect packaging are therefore entering a phase where the challenge is no longer simply demonstrating larger packages or finer interconnects, but ensuring that these complex systems can be manufactured repeatedly, tested effectively, cooled efficiently, and delivered at economically sustainable yields.

The Walls That Must Be Conquered

The fundamental challenge being faced is that several scaling dimensions are increasing simultaneously. AI accelerators require more compute silicon, while growing model sizes and bandwidth requirements are driving the adoption of increasing amounts of HBM. At the same time, chiplet architectures require higher-bandwidth die-to-die interfaces, package dimensions are expanding, and power consumption continues to rise. Each improvement in system capability, therefore, introduces additional complexity somewhere else in the package.

The evolution of large 2.5D packages illustrates the scale of this transition. TSMC's roadmap moves from approximately 5.5 reticle-equivalents for CoWoS in 2026 toward 9.5 reticles in 2027 and approximately 14 reticles in 2028, with the latter expected to support roughly 10 compute dies and 20 HBM stacks. The roadmap extends beyond conventional package dimensions toward much larger wafer-scale configurations.

This is no longer conventional semiconductor packaging in the historical sense. It is increasingly a full system architecture and construction needed to achieve very high performance levels within a complex heterogeneous package.

Unsurprisingly, the economic value concentrated inside this single package also increases dramatically. Instead of assembling one logic die with a relatively simple package, manufacturers may be combining numerous leading-edge compute dies, expensive HBM stacks, silicon or RDL interconnect structures, substrates, and thousands of critical connections. A defect discovered late in this process will represent a much larger financial loss than one discovered during wafer fabrication.

This is exactly where the industry must overcome several of these scaling walls, working together, rather than treating package size, yield, test, power, and thermal engineering as independent problems.

The Different Types Of Advanced Packaging Walls

The first major wall is package-size. Larger packages allow more compute and memory to be integrated, but expanding package dimensions increases warpage, CTE mismatch, routing complexity, and mechanical stress. Large silicon interposers also become more expensive, which is one reason the industry is moving toward architectures that combine silicon only where very high interconnect density is required using RDL or organic structures elsewhere. TSMC's CoWoS-L and CoWoS-R, Intel's EMIB, and Samsung's Cube-E and Cube-R all reflect different implementations of this approach.

The second wall is yield. Chiplets can improve wafer-level economics because smaller dies may yield better than very large monolithic processors, but once many dies are assembled together, package yield becomes dependent on the quality of every component and every assembly operation. Put simply: Assembly yield becomes system yield. As the number of compute dies, HBM stacks, bonding operations, and interconnect structures increases, the economic importance of known-good-die screening and high assembly yield rises rapidly.

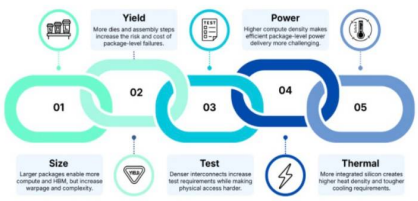

Figure 1: The advanced packaging scaling walls.

The third wall is test and validation. Advanced packaging requires that defects be identified as early as possible because the cost of discovering them increases as additional components are assembled. Wafer sort, known-good-die screening, HBM test, die-to-die interconnect validation, intermediate assembly test, burn-in, final test, and system-level test become part of one continuous test strategy. The source research makes the important observation that test is becoming part of package architecture rather than simply a manufacturing step performed after assembly.

Scaling Wall	What Drives It	Primary Consequence
Package Size	More compute dies, HBM and I/O	Warpage, mechanical stress and interconnect complexity

Yield	assembly steps	Higher economic cost of late-stage defects
Test	More internal and inaccessible interfaces	Greater need for known-good-die and built-in testability
Power	Rising package-level compute density	More complex power delivery and voltage regulation
Thermal	Higher power density and 3D stacking	Cooling becomes a system architecture constraint
Manufacturing	Finer pitches and tighter process tolerances	Difficulty translating demonstrations into high-volume production

Table 1: The drivers and consequences of each wall

Power and thermal limitations also create closely intertwined walls. Adding compute dies and HBM increases both system performance and package-level power density. At the same time, 3D integration reduces interconnect distance by placing dies closer together vertically, but this can make heat extraction significantly more difficult. Power distribution is consequently moving deeper into package architecture, with shorter delivery paths and package-level power-management approaches becoming increasingly important.

These two walls are not independent. Increasing interconnect density may improve bandwidth but make testing harder. Vertical stacking may reduce signaling distance but worsen thermal density. Expanding the package may create room for more HBM but increase warpage and power-delivery complexity. This coupling is what makes advanced packaging fundamentally different from simply making a package larger.

Navigating The Walls Of Advanced Packaging

The scaling walls facing advanced packaging are different from the limitations the industry has dealt with in the past. Package size, yield, testability, power, thermal density, and manufacturing complexity are increasingly connected, which means progress in one area can quickly create pressure somewhere else. Making a package larger is useful only if yield remains manageable, adding more dies is valuable only if they can be tested effectively, and increasing integration density matters only if the resulting system can still be powered and cooled efficiently.

The path forward will therefore depend on balance rather than any single breakthrough. Advanced packaging will have to scale in ways that keep complexity, manufacturability, reliability, and economics aligned as system integration increases. The industry has already shown that packaging can extend semiconductor scaling beyond the limits of a single die. The next challenge is ensuring that these new scaling walls do not simply become the next bottleneck.

In the long run, advanced packaging leadership will be defined less by who can demonstrate the largest package or the finest interconnect, and more by who can turn increasingly complex system integration into a repeatable, reliable, and economically scalable manufacturing platform.

Chetan Arvind Patil is principal engineer for test engineering and customer strategy at Marvell Technology.