Three-Dimensional Integrated Circuits: Technical Foundations, Economic Impact, and Strategic Potential for the Future of Semiconductor Systems

by

Michael Dang

A Thesis submitted to the Faculty

in partial fulfillment of the requirements for the

BACHELOR OF ARTS

Accepted

_____
David Myers, Thesis Advisor

_____
Sami Alshalwi, Second Reader

_____
John B. Weinstein, Provost and Vice President

Bard College at Simon's Rock
Great Barrington, Massachusetts

2025

Table of Contents

Introduction

For over five decades, Moore's Law guided the semiconductor industry by predicting that the number of transistors on a chip would roughly double every two years, leading to exponential improvements in computing power. This relentless 2D transistor scaling has encountered severe challenges in recent years: continued dimensional shrinkage is becoming more difficult and costly, in part because of power-density constraints, and in part because interconnects do not become faster while transistors do. In other words, as transistors switch faster and pack tighter, the global wires connecting them have not kept pace, causing signal delays and power dissipation that limit further gains. These scaling limits – along with quantum-mechanical and material constraints at nanometer scales – signaled an approaching plateau in the benefits of traditional planar integration. The industry has thus been compelled to explore alternative pathways to sustain performance improvement. One promising avenue has been to go vertical: stacking multiple layers of circuits in a three-dimensional configuration. Such 3D integration has been heralded as a new paradigm to extend the spirit of Moore's Law beyond the 2D scaling era. By stacking 2D dies and connecting them in the third dimension, 3D integrated circuits (3D ICs) can pack more devices into a given footprint and shorten critical interconnect lengths, thereby addressing scaling challenges and improving communication speeds between circuit layers. Some researchers even view this as an extension of Moore's Law, achieving greater transistor density per area without solely relying on feature size reduction. In fact, recent commentary suggests that the "new direction for continuing Moore's Law is all about up" – that is, building upwards with 3D architectures.

The rise of 3D integration over the past two decades can be seen as a response to the breakdown of traditional scaling approaches. Around the mid-2000s, Dennard scaling (which had ensured that shrinking transistors also reduced their power density) ceased to hold, leading to a "power wall" in microprocessors. Simply packing more transistors on a 2D die no longer translated to commensurate performance gains due to thermal and power constraints. Additionally, global interconnect delays became a dominant performance bottleneck as clock speeds increased. These trends motivated researchers and industry leaders to investigate 3D IC technologies, which were initially demonstrated in academia and research labs as far back as the 1980s. By the early 2010s, breakthroughs in through-silicon vias (TSVs), wafer bonding, and packaging techniques had matured 3D integration from concept to prototype. The past decade has seen 3D ICs move into production for certain applications – for example, memory stacking in high-bandwidth memory (HBM) and 3D NAND flash memory chips – marking the transition of 3D integration from experimental technology to a practical tool in the semiconductor roadmap. This historical progression sets the stage for a detailed study of 3D ICs and their significance in the future of computing.

Scope of Project

This thesis undertakes a comprehensive investigation into the technological, economic, and strategic dimensions of three-dimensional integrated circuits, a rapidly emerging paradigm in semiconductor design and manufacturing. The project scope encompasses both foundational and applied aspects of 3D integration, aiming to provide an integrated understanding of how vertical chip stacking is reshaping modern electronics.

Technically, the project explores the underlying principles and processes enabling 3D ICs, including through-silicon vias (TSVs), wafer bonding methods, monolithic 3D integration, and thermal management techniques. Emphasis is placed on understanding the practical engineering challenges—such as alignment precision, heat dissipation, and yield degradation—and how modern fabrication and packaging solutions are addressing these issues.

The project also evaluates key 3D IC architectures, comparing fine-grained and coarse-grained partitioning strategies, and investigates design approaches that optimize vertical interconnect density, performance-per-watt, and heterogeneous integration. These analyses span both academic prototypes and commercial implementations, establishing a bridge between research frontiers and industrial practice.

Economically, the thesis evaluates the cost-benefit trade-offs of adopting 3D ICs, focusing on yield implications, manufacturing complexity, chiplet-based economies, and the performance gains that justify higher integration costs. The business case for 3D IC adoption is examined through real-world examples in GPUs, smartphones, FPGAs, and AI accelerators.

Strategically, the scope includes a geopolitical and policy-oriented analysis, focusing on how countries and corporations are investing in 3D IC technologies to secure semiconductor leadership. Special attention is given to government programs like the U.S. CHIPS Act and China's packaging initiatives, highlighting how 3D IC capabilities are becoming pivotal in national security and economic competitiveness.

Finally, this thesis surveys the broader industry ecosystem supporting 3D IC development—such as venture-backed startups, standards organizations like UCIe and IEEE, and

collaborative R&D programs—to show how technical innovation is being matched with ecosystem maturity. The aim is to deliver a multidimensional perspective that informs not only technical research but also strategic planning, investment decisions, and policy formulation related to the future of semiconductors.

Methodology and Sources

To achieve the above objectives, this research employs a combination of literature review and analysis. A wide range of sources has been gathered and synthesized to ensure a thorough and up-to-date understanding of 3D IC technology. These sources include:

Academic literature: Key papers from IEEE and ACM journals and conferences form the technical backbone of the research. Foundational works on 3D IC design, manufacturing, and thermal/mechanical analysis are reviewed, as well as recent publications that report the latest advancements. For example, publications from IEEE International Electron Devices Meeting (IEDM) and IEEE International 3D Systems Integration Conference (3DIC) provide insight into process developments, while design-oriented conferences (like DAC and ISSCC) offer perspective on 3D IC architectures and performance results.

Industry roadmaps and white papers: Important guidance has come from industry-wide collaborative roadmaps, especially the IEEE Heterogeneous Integration Roadmap (HIR) and the International Roadmap for Devices and Systems (IRDS). These documents outline the expected progression of packaging and 3D integration technologies and highlight the challenges anticipated in the next decade. Additionally, technical white papers and reports from leading companies (e.g., Intel's reports on Foveros 3D packaging, TSMC's 3DFabric integration

platform, and reports by research consortia like IMEC) have been consulted to ground this study in practical, state-of-the-art context. Such sources often contain case studies of 3D technology deployment and economic analyses from a commercial perspective.

-Case studies and data: Where possible, real-world examples are used to inform the analysis. This includes documented case studies of products that utilize 3D integration (for instance, AMD's 3D V-Cache technology in CPUs, or the use of HBM memory in GPUs and FPGAs). Market research reports and press releases have provided data points on adoption trends, manufacturing costs, and performance claims, which are critically examined in light of the academic understanding.

By triangulating information from peer-reviewed research, industry roadmaps, and commercial case studies, the methodology ensures a well-rounded perspective. The analysis in this thesis is qualitative in parts (descriptive and comparative) and quantitative in others, drawing on reported experimental results and cost figures from the literature. All sources are cited in IEEE style to allow verification and further reading. In essence, this thesis builds upon the rich body of existing knowledge on 3D ICs while aiming to integrate these insights into a coherent narrative and assessment.

In summary, this thesis Introduction has established the context, motivation, and scope for exploring 3D Integrated Circuits. The subsequent chapters will delve into the technical details, manufacturing techniques, economic considerations, and comparative evaluation of 3D ICs, ultimately providing a holistic understanding of why 3D integration is a pivotal development at this juncture of computing history. By combining insights from engineering and

industry perspectives, the thesis aims to articulate not only how 3D ICs work, but also why they matter for the continuation of progress in electronics in the post-Moore's Law era.

3D Integrated Circuits

A three-dimensional integrated circuit is an electronic chip composed of multiple integrated circuits stacked vertically and interconnected so that they function as a single device. In a typical 3D IC, separate dies (for example, logic, memory, analog, etc.) are thinned and stacked, with vertical connections made through technologies or direct copper-to-copper bonds. By exploiting the vertical dimension (the z-axis), 3D ICs can achieve much higher device density, shorter interconnect lengths, and improved performance and energy efficiency compared to traditional two-dimensional ICs. This approach also enables heterogeneous integration – different technologies can be combined in one stack – which is increasingly important as scaling alone faces limits and system integration becomes a driver of innovation.

3D integration is a broad concept encompassing several levels of packaging and integration. It ranges from 3D packaging (stacking chips or packages at the module level) to true 3D ICs with fine-grained vertical interconnects, as well as intermediate approaches like 2.5D integration using interposers. In general, the term “3D IC” often implies multiple active semiconductor layers within one packaged device, interconnected with high-density vertical links. By reducing the length of interconnect wires and bringing memory closer to logic, for example, 3D ICs promise to overcome performance bottlenecks of 2D chips such as limited bandwidth and long global interconnect delays. At the same time, 3D IC technology introduces new challenges in design, thermal management, manufacturing, and testing, which require significant research and development effort. This part of the Thesis presents a comprehensive literature study of

3D ICs, covering their history, architecture, thermal and manufacturing issues, testing and reliability, applications, and recent advances.

Motivation for 3D ICs – Technical and Economic Drivers

The motivation for studying and developing 3D integrated circuits is twofold, encompassing both technical imperatives and business/economic pressures. On the technical side, 3D integration directly addresses several critical limitations of planar circuits:

Interconnect Delay and Bandwidth

In large 2D chips, signals must traverse long metal interconnects, incurring latency and power loss. Stacking chips in 3D shortens the average interconnect length, which can speed up communication between functional blocks on different layers. By providing abundant vertical connections (through TSVs or bonding interfaces), 3D ICs enable wider bandwidth buses between, for instance, a processor and a memory layer, vastly increasing off-chip memory communication throughput. This helps alleviate the notorious “memory wall” problem that plagues modern processors by bringing memory physically closer to logic. A prime example is logic-on-memory stacking: a processor die bonded with a DRAM cache die can achieve bandwidth far beyond what is possible with conventional packaging.

Power and Energy Efficiency

Shorter interconnects and on-chip vertical vias do not only improve speed but also reduce power consumption. Signals that remain on-chip avoid the large power overhead of driving off-chip I/O; studies have shown that keeping communication on-chip can reduce signal power by an order of magnitude (10× to 100×). Furthermore, with 3D integration, the reduced wire length and capacitance translate to lower switching energy per interconnect. In aggregate, a well-designed 3D IC can operate more energy-efficiently than an equivalent 2D implementation, which is crucial as power has become the limiting factor in many designs.

Heterogeneous Integration and Functionality

3D ICs allow different circuit layers to be manufactured in disparate process technologies optimized for their purpose. For example, analog/RF circuits, high-density memory, and logic processors could be fabricated on separate wafers using the best-suited process for each, then vertically integrated into one composite chip. This capability means that "circuit layers can be built with different processes, or even on different types of wafers," permitting far greater optimization and the combination of components that would be incompatible if made on a single substrate. The result is a single 3D IC that might include, say, a CMOS logic layer, a III-V compound semiconductor layer for high-speed I/O, and a memory layer – achieving in one stack what traditionally would require multiple chips. Such heterogeneous 3D integration not only improves performance but also opens new possibilities for system architecture that are not feasible in 2D.

Compact Form Factor

By building “up” instead of “out,” 3D chips can dramatically reduce the board area or package footprint required for a given functionality. This is especially important for mobile and wearable devices, IoT sensors, and other form-factor-constrained systems. For instance, stacking memory arrays vertically enables storage devices with much higher capacity per unit area than planar flash memory. From an architectural perspective, increased integration density can be seen to continue functional scaling even when lateral scaling slows – effectively achieving “more transistors per footprint” to extend Moore’s Law in terms of device density.

Historical Development of 3D IC Technology

The vision of building circuits in three dimensions emerged soon after the advent of the integrated circuit itself. As early as the 1960s, engineers had proposed stacking active devices to increase density. In 1964, researchers at Texas Instruments described the concept of a vertically integrated MOS IC, and in 1969, a team at NEC in Japan proposed a 3D stacked memory chip concept. Early implementations of multi-chip stacking were seen in specialized military or aerospace electronics using 3D packaging. However, these were package-level integrations, not monolithic 3D ICs.

Modern 3D IC research really took off in the 1980s, notably in Japan. In 1981, a national R&D project in Japan called “Three-Dimensional Circuit Element” began exploring 3D integration technologies. Two main approaches were investigated: recrystallization, where a silicon layer is deposited and recrystallized on top of an

existing wafer to form devices, and wafer bonding, where separately fabricated wafers are bonded together. By 1983, researchers at Fujitsu had successfully fabricated a simple 3D CMOS inverter stack using laser beam recrystallization – essentially placing a transistor directly above another transistor with an insulating layer in between. In 1984, the same group demonstrated a 3D CMOS gate array on a dual-layer silicon-on-insulator structure. Around the same time, teams at Mitsubishi and NEC pursued similar concepts. In 1987, Mitsubishi Electric built a three-layer 3D image signal processor with photodiodes, analog-to-digital converters, and logic ALUs on different layers. By the end of the 1980s, NEC had fabricated a four-layer 3D IC using laser recrystallization. These Japanese efforts were pioneering, establishing basic 3D chip architectures, and proving that stacking active devices was feasible, though the methods were limited in device quality.

In addition to recrystallization approaches, Japan also led early development of wafer bonding with vertical interconnects. The concept of bonding thinned device wafers was explored in the 1980s and reached a milestone in 1990 when an NEC team led by Hayashi demonstrated a bonded two-layer NMOS device with inter-wafer connections. They proposed that this technique could be extended to many layers by thinning and bonding wafers in sequence. In fact, the first through-silicon via structures were conceived in Japan during this era: Hitachi filed a patent in 1983, Fujitsu in 1984, and by 1986 Fujitsu described a chip stacking method using vertical vias through silicon. Prof. Mitsumasa Koyanagi's group at Tohoku University was especially instrumental: in 1989 they developed wafer-to-wafer bonding with TSVs and demonstrated a 3D LSI chip.

Throughout the 1990s and early 2000s, Koyanagi's group continued to push 3D integration – by 2000 they had a three-layer memory chip, by 2001 a three-layer "artificial retina" sensor chip, and by 2005 a ten-layer memory stack, all using TSV-based wafer bonding. These achievements established the fundamental technology of TSV interconnects and showed that quite complex, multi-layer devices were possible.

In the United States, 3D IC work accelerated in the late 1990s. At MIT, Fan, Rahman, Reif and colleagues developed a copper-to-copper wafer bonding method in 1999, achieving extremely fine-pitch vertical connections without solder. This copper bonding technique would later become the foundation of modern hybrid bonding processes. In the early 2000s, the U.S. DARPA (Defense Advanced Research Projects Agency) initiated programs to push 3D IC technology, providing funding that spurred academic and commercial prototypes. One of the first fabless startups focused on 3D ICs, Tezzaron Semiconductor, built working 3D chips by 2004 using a "via-first" TSV process and wafer bonding. Tezzaron demonstrated a stacked memory and an 8051 microcontroller with vertically integrated memory, showing significantly improved speed and power consumption over its 2D equivalent. Around the same time (2004), Intel revealed a research prototype of a 3D Pentium 4 processor, implemented with two active tiers bonded face-to-face. In that experimental 3D Pentium, functional blocks were split between two dies to reduce pipeline stages and long wires, resulting in ~15% performance improvement and 15% power reduction compared to a conventional Pentium 4, thanks to eliminated repeaters and shorter interconnects.

By the late 2000s, more complex 3D chips were prototyped. Intel's "Teraflops" 80-core research chip in 2007 stacked a logic layer with a high-bandwidth memory layer using TSVs, achieving an unprecedented 1 TB/s memory bandwidth for the 80 cores while drastically cutting power for off-chip I/O. Academic researchers also joined the effort: for example, in 2008, University of Rochester demonstrated a 3D microprocessor operating at 1.4 GHz, exploring how to route clock and data vertically through layers. In 2012, two well-known academic 3D multi-core prototypes were presented: 3D-MAPS, a 64-core processor from Georgia Tech with two logic layers, and Centip3De, a 64-core ARM Cortex-M3 based design from University of Michigan operating at near-threshold voltages on a tiered stack. These prototypes, built using GlobalFoundries 130 nm processes and Tezzaron's FaStack bonding, were important in validating 3D IC advantages in a realistic computing and also in identifying issues like heat and toolflow limitations.

Meanwhile, industry was bringing the first 3D-stacked products to market. An early commercial utilization came in 2004 when Sony's PlayStation Portable incorporated a 3D stacked DRAM memory in its graphics engine chip package. This was a two-die stack (logic+DRAM) in a system-in-package, and Toshiba dubbed it "semi-embedded DRAM" – an early example of integrating memory on logic to boost bandwidth for a handheld device. In 2007, Toshiba further commercialized an eight-layer stacked NAND flash memory chip (16 GB) by vertically integrating eight 2 GB NAND dies for data storage. Hynix soon followed with a 24-layer flash stack (16 GB) using wafer bonding. Flash memory, used in memory cards and SSDs, quickly adopted die-

stacking to achieve higher densities: by 2010 Toshiba was selling 16-die flash stacks (128 GB). Stacking in this context was at the package level, but it demonstrated the viability of mass-producing 3D memory. DRAM vendors also embraced 3D techniques in the early 2010s: Elpida developed a 4-die stacked 8 GB DDR3 SDRAM module in 2009. By 2011, both Hynix and Samsung had unveiled TSV-based DDR3 DRAM stacks. That same year, industry consortiums led by Micron and Samsung announced the Hybrid Memory Cube (HMC), which uses a logic layer and several DRAM layers in a single stack interconnected by TSVs. These developments culminated in the introduction of High Bandwidth Memory (HBM) around 2013. The first HBM stacks (4-high DRAM dies on a base logic die) were manufactured by SK Hynix in 2013. HBM technology was soon adopted in GPUs and high-performance computing; by 2016 Samsung was in mass production of second-generation HBM2 stacks up to 8 GB each.

Today, 3D integration has become mainstream in certain domains, especially memory. Virtually all high-density NAND flash memory is produced as 3D NAND, which is a monolithic stacking of memory cells. High-bandwidth memory and 3D DRAM are employed in cutting-edge GPUs, FPGAs, and network processors to overcome memory bandwidth bottlenecks. Processor manufacturers have also begun to integrate logic-on-logic stacks: for example, AMD's Zen 3 and Zen 4 CPUs use "3D V-Cache" technology, where an extra SRAM cache die is bonded atop the CPU die to provide a larger L3 cache in 3D. Intel's Foveros technology similarly enables stacking a small high-density logic die on a base die used in products like Intel Lakefield and upcoming Meteor Lake chips. These recent breakthroughs mark the transition of 3D ICs

from experimental to commercial use in processors, and ongoing advancements like hybrid bonding with <10 μm pitch, and monolithic 3D at transistor-level continue to push the boundaries of integration.

3D IC Architectures and Design Approaches

Designing a 3D IC requires deciding how to partition and stack the system's components across multiple layers. There are different architectural design styles, generally categorized by the granularity of 3D partitioning. At one extreme, fine-grained 3D integration distributes individual logic gates (or small standard-cell groups) across layers. At the other extreme, coarse grained 3D integration assigns entire functional blocks or IP cores to different dies in the stack. Early academic studies suggested that fine-grained 3D integration could yield the greatest wirelength reduction and density gains, since each logic block could potentially be split between layers to minimize interconnect. However, fine-grained 3D design faces major practical challenges: it requires an enormous number of TSV connections between layers, complicates testing, and demands new EDA tools for 3D placement and routing which are not yet fully mature. In fact, splitting logic gates across dies can hurt yield — a single defective die could disable the entire module — and exacerbate process variation issues across layers. For these reasons, coarse-grained 3D architectures are generally preferred in practice.

In coarse-grained 3D ICs, each die contains one or more large functional blocks (e.g. a processor core, a cache bank, an analog front-end, etc.), and only a relatively limited number of global signals cross between dies via TSVs. This style reduces the

required TSV count and allows most IP blocks to remain intact (designed and verified with conventional 2D tools), easing integration. In fact, one can often use existing 2D-designed blocks on each layer and just handle the 3D connections and thermal management at a higher level, which is far more convenient than redesigning everything for 3D. Many modern 3D chips follow the coarse-grained approach – for example, a stacked processor and memory is a coarse 3D partitioning. This approach also aligns with heterogeneous integration: different dies can be implemented in different technologies (one in a high-performance logic process, another in a dense memory process, etc.) and then integrated vertically.

Regardless of granularity, a key architectural consideration is the interconnect network between layers. 3D ICs introduce vertical interconnects (TSVs or micro-bumps) that can be treated as additional wiring resources in the chip design. Often a Network-on-Chip (NoC) or bus is extended into the third dimension. For instance, in a 3D multi-core processor, each die might have its own local interconnect fabric, with vertical links connecting routers or buses between dies. The vertical communication latency is very small (TSV connections are short, on the order of tens of microns), so architects can leverage this to, say, place a cache directly above a core and connect them with dense TSV pillars, achieving bandwidth that would be impossible with off-chip links. Banerjee et al. (2001) proposed a seminal 3D IC architecture where global interconnect bottlenecks in a 2D chip are relieved by splitting the logic across two layers and inserting plentiful vertical vias between them. Their study showed that critical paths could be shortened and heterogeneous functions (like analog, memory, logic) integrated on separate planes,

significantly improving performance of a system-on-chip. The general finding across many such studies is that 3D architectures allow new trade-offs: one can reduce clock cycles or pipeline depth (since signals don't have to travel as far), use wider buses between units (since vertical via density can be high), and even implement novel structures like tiered caches or memory-on-processor architectures that are not possible in 2D.

Another architectural taxonomy for 3D ICs is by the level of integration. As noted, "3D IC" often implies TSV-based die stacking, but we also have 2.5D integration, where multiple dies are placed side by side on a silicon interposer. In 2.5D, the interposer provides a passive wiring plane with TSVs to an underlying package substrate, enabling high-density connections between chips (e.g., used in high-end FPGAs and GPUs). 2.5D is sometimes considered a subset of 3D integration since it uses TSVs and advanced packaging, though the dies themselves are not stacked. A real example is Xilinx's Virtex-7 2000T FPGA (2011), which mounted four FPGA slices on a silicon interposer in one package – effectively creating a larger "virtual" chip – something not achievable as a single die due to yield and size limits. That product demonstrated that combining multiple smaller dies on an interposer (an approach often called chiplet integration today) can deliver the benefits of a very large chip, illustrating one end of the 3D integration spectrum. At the other end is monolithic 3D integration, where multiple layers of transistors are fabricated sequentially on one wafer. Monolithic 3D (also called sequential 3D IC) can achieve extremely fine vertical connectivity – essentially at the vias-between-transistors level – far beyond the density of TSVs. For example, researchers have

demonstrated CMOS circuits with a second layer of transistors built directly atop a first layer, connected by nano-scale vertical vias similar in size to regular interconnect vias. Monolithic 3D promises the ultimate integration density and has been shown in experimental devices (often using processes like wafer bonding of thin transistor layers or laser crystallization for upper layers), but it is not yet commercially deployed due to process complexity and thermal constraints. Companies like IBM have investigated "layer transfer" techniques to achieve something close to monolithic integration – transferring a finished device layer onto another wafer – and managed sub-micron alignment and very high via densities in experimental 3D chips.

Current 3D IC designs tend to use the coarse-grained approach (each layer handling different large functions) and rely on TSV-based interconnects or an interposer, because this is manageable with existing design flows. This still provides significant benefits: for instance, memory can be placed directly above logic to vastly increase bandwidth, or analog/RF circuits can reside on a separate die to avoid digital substrate noise. As EDA tools and manufacturing techniques improve, finer integration might become viable. In any case, architects must also consider issues like heat distribution and power delivery when designing a 3D IC, which leads to the next topic – thermal management.

Thermal Management in 3D ICs

Thermal issues are one of the most critical challenges for 3D ICs. By stacking active layers, the power density (power per unit footprint area) increases, and it becomes harder for heat to escape from the interior of the stack. In a 2D chip, each transistor is relatively close to the surface (which is attached to a heat spreader or heat sink). In a 3D IC, a transistor buried in the middle of a stack has other active layers above and below it acting as insulation. As a result, hotspots can build up deep in the stack and raise the temperature significantly if not properly managed. The basic problem is that electrical proximity correlates with thermal proximity: grouping circuits closer in space means their heat is also concentrated. Without adequate cooling, a 3D IC could overheat even if each layer individually would be cool in a 2D implementation.

A straightforward way to mitigate thermal issues is clever architectural thermal planning: for example, designers can place high-power blocks on the top die (closest to the heat sink) or spread power-hungry units across different layers, so they are not all directly on top of each other. In the Intel 3D Pentium 4 prototype, designers split large, hot units across the two dies and rearranged them to limit thermal coupling, which helped keep the chip within thermal limits. Another technique is inserting thermal vias – these are dummy or dedicated TSVs filled with thermally conductive material that do not carry signals but help conduct heat out of the stack. Thermal TSVs can be sprinkled under hotspots to provide heat removal pathways.

Despite these design measures, advanced cooling techniques are often needed for high-power 3D ICs. One promising solution researched extensively in the last decade is microfluidic cooling: integrating tiny liquid coolant channels or heat sinks directly into the stack. In fact, DARPA launched a program called “Intra/Interchip Enhanced Cooling” to investigate using liquid cooling within 3D chip stacks. Researchers have demonstrated microchannels etched in silicon through which coolant is pumped to carry away heat from each layer. Reviews of 3D IC cooling categorize various approaches: some embed microchannels between layers, others use piped coolant through TSV-like micro conduits, and some combine conventional heat sinks with embedded cooling structures. Microfluidic cooling has shown the potential to remove heat densities on the order of hundreds of W/cm^2 from 3D stacks, far beyond what passive heat spreading can do, effectively extending the thermal envelope of 3D ICs. Of course, integrating liquid cooling adds complexity in manufacturing and design.

Apart from liquid cooling, other thermal management research includes use of high-conductivity interface materials, heat spreaders within the package. In practice, today’s commercial 3D ICs tend to have moderate power per die and are often mounted on an interposer or substrate that can distribute heat. But for logic-heavy 3D chips, thermal management remains a gating issue. Software and design-time thermal analysis is also crucial – 3D IC designers use thermal simulation tools to predict hotspots and ensure that even worst-case power scenarios won’t exceed temperature limits. To summarize, the thermal challenge of 3D ICs is inherent to the technology: stacking multiplies the heat density. Solutions like careful floor planning, thermal vias, and advanced cooling

(microchannels, two-phase cooling, etc.) are active research areas aimed at enabling high-performance 3D chips without overheating. This is a key area where 3D IC design must innovate beyond traditional 2D cooling techniques.

Manufacturing Techniques for 3D ICs

Building a 3D IC is considerably more complex than fabricating a 2D chip, as it involves additional process steps to create and align the vertical connections and to bond multiple wafers or dies. Several distinct manufacturing approaches have been developed:

Through-Silicon Vias (TSVs)

To make TSVs, one must thin the wafer (so the vias are short), etch holes (often ~5–10 μm diameter for many applications) from one side, insulate and fill them with copper, and land them on pads on the other side. TSVs are large relative to transistors – for example, a 10 μm×10 μm TSV might take the area of tens or hundreds of logic gates– so they impose a penalty on area and routing. The manufacturing of TSVs can be done via-first, via-middle, or via-last:

- Via-first means the TSV is made early in the process, before the front-end (transistor) fabrication is complete.
- Via-middle is after front-end but before the back end (metal layers) are complete.
- Via-last inserts TSVs after the wafer is largely finished (through the completed die).

Each approach has trade-offs in terms of how the TSV integrates with the devices and metallization. Via-first TSVs consume device area (since they are present during transistor formation), whereas via-last TSVs must drill through finished structures and thus are constrained in size and placement. Regardless of method, TSVs also require keep-out zones where no active devices can be too close (to avoid stress and interference), which further increases their effective footprint. A careful balance is required in deciding the number and size of TSVs: enough to meet bandwidth and power needs, but not so many that they bloat the die area or introduce too much capacitance.

Wafer Bonding and Stacking

Once dies or wafers have TSVs (or at least pads for vertical connections), they need to be physically bonded together in alignment. There are a few bonding techniques:

- Die-to-Die (D2D) bonding: Two completed dies are aligned and bonded.
- Die-to-Wafer (D2W) bonding: A diced die is bonded onto a still-intact wafer that has multiple chips on it.
- Wafer-to-Wafer (W2W) bonding: Two full wafers (each containing many chips) are aligned and bonded, and then later diced into stacked chips.

Wafer-to-wafer bonding can be more efficient (bond many dies at once), but it requires very high yield and alignment accuracy across the entire wafer – any bad die will make the whole stacked pair bad, so typically both wafers must be defect-free in corresponding die positions or have some redundancy. The bonding itself can be done by fusion bonding (direct bonding of insulator or silicon surfaces, often followed by via formation for connectivity) or metal bonding. Copper-copper bonding is a popular

approach for 3D ICs: here, copper pads or posts on each die are pressed together under heat/pressure to form a solid bond. This was pioneered at MIT as mentioned (Cu-Cu wafer bonding) and provides very high electrical and thermal conductivity. Modern processes use hybrid bonding, where a dielectric (oxide) bond provides mechanical strength and embedded copper regions bond to provide electrical connection. Hybrid bonding can achieve extremely fine pitches (on the order of a few microns or less), far tighter than traditional micro bumps. In fact, technologies like TSMC's SoIC and Intel's Foveros Direct are essentially advanced copper-copper bonding allowing direct die stacking without bumps, enabling near-monolithic connection density in production. Another bonding method is solder-based microbumps: small solder balls (say 20–50 μm) are formed on TSV landing pads and reflowed to join dies. Microbumps have been widely used, but their pitch is limited by solder bump size, and they introduce additional resistance. The trend is toward bump-less bonding (oxide/copper hybrid).

Precise alignment is crucial during bonding. Modern bonders can align wafers or dies to within a fraction of a micron. IBM's research in the 2000s achieved wafer-to-wafer alignment under 1 μm, enabling TSV densities over 10^8 vias/cm² in a layer-transfer 3D process. Such alignment fidelity allows extremely dense vertical interconnect — an example reported by IBM was contacts with aspect ratio 6:1 to 11:1 between two stacked device layers, only 2 μm apart vertically. These figures approach the regime of monolithic 3D, showing that with careful process engineering, 3D IC bonding can achieve very high inter-layer via density and very short connections.

Thinning

To make effective 3D stacks, chips often need to be thinned after bonding (if bonded face-to-face) or before bonding (if face-to-back). Thinning reduces the thermal resistance and shortens TSV lengths. Typical thinned die thicknesses are 50–100 μm for many TSV applications and can be as low as a few tens of microns or even <10 μm in some experimental stacks. Thinning is done by back-grinding and chemical-mechanical polishing, and the thin wafer may be supported on a temporary carrier during processing.

Monolithic Integration Processes

In contrast to stacking completed die, monolithic 3D integration uses semiconductor process flows to build transistors layer by layer. This might involve depositing a silicon layer on top of an existing chip and using processes like recrystallization or epitaxy to create a device-quality layer, then repeating transistor fabrication on that new layer. Another approach is the layer transfer: fabricate a layer of devices on one wafer, then peel and bond that thin layer onto another wafer. After bonding, inter-layer vias are formed to connect the devices. These processes require compatibility with thermal budgets (the upper devices must be made at temperatures that won't damage the lower devices) and planarity for bonding. While not yet commercial, these techniques are being refined. For instance, a recent demonstration of monolithic 3D stacked CMOS showed vertically integrated inverter circuits where the upper device layer was formed at low temperature with only minor performance penalty.

Manufacturing 3D ICs also brings yield challenges.

Each additional process step (etching vias, bonding, thinning) can introduce defects. If two dice are bonded, a defect in either can ruin the pair, so the effective yield is the product of yields of each layer. This has driven the development of Known Good Die (KGD) strategies – testing dies before stacking to ensure only good ones are integrated. Some 3D flows use partial stacking and then test the stack before adding more layers to avoid wasting too many good dies with one bad. We will discuss testing approaches in the next section. Overall, from a manufacturing standpoint, cost and complexity were initially high barriers for 3D ICs, but advances have streamlined many steps. Analyses have shown that if the key cost drivers are managed, 3D ICs can be economically viable. In fact, the cost can be justified when 3D integration replaces expensive large-die monolithic designs or yields a substantial performance/power benefit that justifies a premium. Today, foundries offer TSV processes and packaging houses offer stacking services, indicating that 3D IC manufacturing has matured significantly, at least for certain product types (memory, etc.). Ongoing research continues to improve techniques like finer-pitch TSVs, better alignment, and more reliable bonding processes, which will further enhance the manufacturability of 3D ICs.

Testing and Reliability of 3D ICs

Testing 3D ICs is considerably more difficult than testing conventional 2D chips. In a stacked 3D device, once dies are bonded together, it becomes challenging to access internal signals of the middle layers from the outside world. A fundamental principle in

test is to perform pre-bond testing – i.e., test each individual die (while it is still separate) to ensure it is a Known Good Die before assembly. Pre-bond testing maximizes the yield of the final stack by not assembling bad dies together. However, pre-bond testing requires that each die has some temporary probe access to its I/Os. For dies that will be internally sandwiched, their primary I/Os are meant to connect to other dies (through micro-bumps or TSVs) which are not easily accessible with probe needles. Specialized probe pads or breakable connection points can be designed to facilitate pre-bond test, but this adds overhead. Researchers have proposed solutions like built-in self-test circuits and dummy probe pads or e-fuses that can connect internal TSV networks to probe-able pads during test. IEEE tackled this problem by developing a design-for-test standard for 3D ICs: IEEE Std 1838-2019, which defines a Test Access Architecture for 3D stacks. The IEEE 1838 standard basically provides a blueprint for adding a "die wrapper" interface on each die so that when dies are stacked, their test chains can be interconnected and accessed from the top or bottom of the stack. This allows, for instance, a JTAG or boundary scan chain to snake through all dies, so that testing can be done post-assembly, and each die in the stack can be isolated or tested in tandem. With such standards, automatic test equipment can treat the 3D stack in a structured way, improving test coverage.

Still, some testing is ideally done at multiple stages: pre-bond, mid-bond, and post-bond. Pre-bond (each die separately) catches defects early. Mid-bond testing might involve testing partial stacks (for example, after bonding two dies, test that interface before adding a third). Post-bond (final) testing tests the fully assembled 3D IC like a normal chip to catch any integration-induced faults. Key challenges remain in each stage:

for example, probing micro-bumps on a very thin die is mechanically tricky. One approach in industry has been to include temporary probe pads on the wafer which are only used for test and then discarded (or limited probe of TSV tips if accessible). The IEEE 1838 standard and similar DfT architectures help by allowing stacked-die test using minimal test pins – essentially, the dies collaborate to present a unified test interface to external equipment.

Another aspect is internal test isolation. If a module is split across dies, traditional test patterns might not detect faults easily because part of the logic lies in one die and the rest in another. For timing-critical paths that traverse dies, new test strategies are needed to ensure faults (like TSV defects or inter-die timing violations) are caught. Research and industrial efforts have extended scan testing, built-in self-test (BIST), and at-speed test methods to the 3D context. For instance, 3D scan chains can be configured that include TSVs as scan elements to test them. Memory-on-logic stacks might use BIST engines on one die to test memory on another, etc.

Reliability is closely tied to testing – many reliability concerns must be designed and tested for. One major reliability issue in 3D ICs is thermomechanical stress. TSVs, being copper in silicon, have a different coefficient of thermal expansion than the silicon matrix. During thermal cycles (from fabrication through operation), the mismatch can induce stress in the silicon around TSVs, potentially shifting transistor parameters or even causing microcracks if not managed. Likewise, bonding different materials in a stack creates internal stresses. A 3D stack has a complex mechanical profile: multiple

thin silicon layers, bonding interfaces, and copper interconnects all expand and contract with temperature, which can lead to warpage or strain. This can impact reliability through phenomena like TSV delamination or cracking, die warpage breaking bonds, or accelerated electromigration in the TSVs and vertical interconnects due to local thermal hotspots. Designing keep-out zones (where no sensitive transistors sit too close to a TSV) and using dummy TSV insertion to balance stress are common practices to mitigate stress effects. Additionally, thermal cycling tests and stress simulations are important to ensure reliability over the product's lifetime.

Another reliability factor is power delivery and noise in 3D stacks. If each layer draws current, the power must be delivered either through TSVs or from one layer to the next. This can introduce larger IR drops or LDI/DT noise coupling between layers. Techniques like having separate power TSVs, or even using fluid cooling channels to deliver power (as IBM has explored with a liquid that carries both power and cooling) have been considered. Signal integrity of TSVs (crosstalk, etc.) and the impact of processing (for example, TSV etching can damage nearby devices) also need to be accounted for in design and verification.

In summary, testing and reliability for 3D ICs demand new strategies beyond conventional IC test. The need to test dies in isolation and then as a unit is critical for yield. The introduction of IEEE 1838-2019 provides a standardized way to design testability into 3D ICs, indicating the maturation of this field. Reliability challenges like thermomechanical stress require co-design of the mechanical and electrical aspects of the

chip. Despite these challenges, successful 3D products have shown that with careful design for test and reliability 3D ICs can meet commercial quality and reliability standards.

The Business, Economics, and Strategic Importance of 3D Integrated Circuits

This chapter provides a comprehensive overview of the business and economic considerations of 3D IC adoption and its strategic importance. We will examine the cost-benefit trade-offs (such as performance-per-watt gains versus yield and manufacturing challenges), highlight major application domains (data centers, smartphones, and AI hardware), and explore how 3D integration figures into national security and semiconductor sovereignty efforts. We will also survey current investment trends, including venture funding and corporate R&D, surrounding 3D IC technologies. Real-world examples from industry leaders (TSMC, Intel, AMD, NVIDIA, Samsung, and others) and government initiatives (U.S. CHIPS Act and Chinese programs) will illustrate these points.

Cost-Benefit Analysis of Adopting 3D ICs

Adopting 3D IC technology involves balancing significant benefits against the costs and challenges introduced by this approach. From a business perspective, companies must weigh improvements in performance and energy efficiency and potential system-level cost savings against the higher manufacturing complexity and possible yield

issues of 3D integration. Here we analyze these trade-offs in terms of performance-per-watt advantages, yield and defect considerations, and manufacturing complexity/cost.

Performance-Per-Watt and Density Benefits

One of the strongest arguments for 3D ICs is the improvement in performance-per-watt and in overall package density that vertical integration can provide. By stacking chips, the distance that signals need to travel between functional units is greatly reduced, which lowers communication latency and power consumption per bit of data transferred. A well-designed 3D IC can thus achieve higher performance at the same power budget (or lower power for a given performance) compared to a traditional 2D system. Industry experts note that 3D integration offers increased performance, lower power consumption, and miniaturization of electronics, which is attracting interest for applications ranging from mobile devices to high-end AI and data center. In practical terms, stacking a high-bandwidth memory (HBM) die directly on top of or beside a logic die (CPU/GPU) in one package eliminates the power-hungry off-chip memory interface; this delivers orders-of-magnitude higher memory bandwidth at a fraction of the energy per data transfer. For example, the use of 3D-stacked HBM memory in modern GPUs and AI accelerators provides memory bandwidths in the terabytes per second while using significantly less power per bit than traditional DDR/GDDR memory mounted on a board. Moreover, 3D ICs enable greater functional density – more transistors or bits per unit area – which is valuable when space is constrained (such as in smartphones or wearables). By building vertically, 3D ICs can fit the same functionality in a smaller footprint or add new capabilities without increasing package size. This vertical densification is analogous to

constructing a high-rise building instead of many single-story buildings to pack more capacity in the same land area.

Heterogeneous Integration and System Cost Advantages

Another benefit of 3D ICs is the ability to integrate heterogeneous technologies and reuse modular "chiplets," which can bring cost and flexibility advantages. Different components of a system (logic, memory, analog, RF, etc.) can be manufactured on the process technology that best suits them and then bonded together in a 3D package. This means a complex system does not have to be implemented on a single large, cutting-edge monolithic die. Instead, designers can disaggregate a system-on-chip (SoC) into smaller chiplets, each built on an optimized process node (e.g., some on advanced nodes for dense logic, others on older nodes for I/O or analog that scale poorly). These chiplets can then be interconnected within one package. Such chiplet-based heterogeneous integration delivers greater performance at a reduced cost and higher yield compared to a traditional monolithic SoC, with only a minor area.

The cost reduction comes from several factors. First, splitting a design into multiple smaller dies can improve overall yield: a smaller die has a lower probability of containing a manufacturing defect, so more of them are functional. Instead of one large die that might be lost to a single defect, multiple small dies increase the chances that most of the system's area is defect-free. In fact, dividing a large chip into chiplets can yield better than a single huge die, which reduces waste and lowers effective cost per good chip. Second, not every part of a system needs the most expensive silicon technology –

for example, high-precision analog blocks or I/O interfaces might see little benefit from a 5 nm process and can be built on a cheaper 28 nm process. 3D integration allows such components to be implemented separately and then integrated, mixing different process nodes in one package. This heterogeneous approach can reduce risk and cost as well; proven designs/ existing chips can be repurposed and stacked rather than redesigning everything into one new chip, which leverages prior investment and avoids the expense of a full SoC re-design. Overall, from a business standpoint, 3D chiplets and heterogeneous integration can shorten design cycles and cut development costs by enabling IP reuse and modularity, while also controlling manufacturing costs by using each technology node appropriately. These benefits, however, are realized only if the integration technology is sufficiently developed and does not introduce prohibitive overhead.

Yield Concerns and Debug/Test Challenges

The flip side of stacking multiple chips is the impact on yield and the complexity of testing. Manufacturing yield – the fraction of produced devices that are functional – is a critical determinant of cost. 3D IC approaches can both help and hurt yields, depending on how they are implemented. On one hand, as noted above, breaking a large design into smaller dies can improve yields for each die, and known-good dies can be selected and combined, avoiding the need to discard a large monolith if one part is bad. On the other hand, certain 3D integration methods (notably wafer-to-wafer bonding) compound defect rates: if you directly bond whole wafers of dies, the combined yield is roughly the product of the yields of each layer, meaning the more layers you stack, the more overall

yield can suffer. A classic concern is that if any 1 of N chips in a 3D stack is defective, the entire 3D IC becomes defective, so stacking without pre-testing can drastically reduce final yield.

This is why many 3D IC assembly flows use die-to-die or die-to-wafer bonding with known-good die selection or incorporate built-in redundancy and repair. Each additional processing step or bonding interface in a 3D build is also a potential site for defects or misalignment. Therefore, sophisticated testing strategies are needed to identify bad dies before stacking and to verify 3D interconnects after stacking. Researchers emphasize that to achieve high overall yield and reasonable cost, it is essential to test independent dies separately before integration (known-good die approach). Testing a 3D IC is more complicated than testing a single 2D chip because once stacked, accessing internal die nodes is difficult; special test structures or design-for-test techniques must be employed for the vertical interconnects and tiers. In summary, the yield issue is a major consideration: companies must invest in advanced test and repair methods to ensure that stacking does not wipe out the economic gains. If done poorly, a 3D layout might paradoxically yield worse than an equivalent 2D design, so careful partitioning and integration strategy are key to reaping the benefits.

Manufacturing Complexity and Thermal Management

3D IC integration is technically complex, which translates to higher manufacturing costs and new engineering challenges that must be managed. Building a reliable 3D IC involves additional process steps such as TSV formation (drilling and filling micron-scale vias through silicon), wafer thinning (to make stacked dies thin

enough), high-precision alignment and bonding of multiple dies, and often specialized cooling solutions. These steps require capital-intensive equipment and meticulous process control, driving up fabrication and packaging costs. Only a few advanced foundries and OSATs (Outsourced Semiconductor Assembly and Test companies) have the capabilities to do fine-pitch 3D stacking, which can create supply constraints and higher prices for 3D packaging services.

Designing a 3D IC is also more complex: engineers must consider the third dimension in all stages of design, from architecture to physical implementation and thermal analysis. Traditional optimization metrics of power, performance, and area (PPA) become power, performance, and area-density in a 3D context – effectively optimizing per cubic millimeter instead of per square. This holistic co-design requirement means that chip design teams often need new electronic design automation (EDA) tools and expertise to plan power delivery and signal routing in 3D, as well as to co-design the package and dies together. For example, the power delivery network in a 3D stack must be carefully engineered so that upper dies receive stable power through TSVs or inter-die interconnect from lower layers.

Thermal management is arguably the biggest hurdle in 3D ICs: stacking active devices increases local power density and makes cooling more difficult, since interior layers are insulated by silicon and dielectrics with poor thermal conductivity. If one die runs hot, it can heat up adjacent dies, potentially creating thermal throttling issues or even

thermal-induced failures if not managed. Thus, additional costs may be incurred for advanced cooling solutions and for extensive thermal simulations during design.

Finally, the assembly yield and reliability of 3D connections (TSVs, micro-bumps, or hybrid bonds) are concerns – voids or misalignments in bonding can cause failures that are hard to detect and repair. All these complexities mean 3D ICs often have a higher upfront manufacturing cost per unit and higher non-recurring engineering (NRE) costs for design and verification. However, these costs may be justified by the substantial performance gains and potential system-level cost savings discussed earlier. In niche high-performance markets, the cost per additional FLOP or per watt saved may make 3D integration a compelling investment despite the complexity. Industries and governments are actively funding R&D to streamline these processes and bring down the cost of 3D manufacturing, in hopes of eventually making 3D ICs more routine.

In summary, the cost-benefit equation for 3D ICs must account for significant benefits – notably improved performance-per-watt, higher density, and heterogeneous integration enabling better yield and modularity – against the costs in yield complexity, manufacturing difficulty, and required engineering innovation. Many early 3D IC deployments have been in applications willing to pay a premium for performance or form factor, but as the technology matures, the expectation is that costs will decrease. Industry reports already indicate that chiplet-based designs can achieve greater performance at reduced cost and with higher yield than comparable monolithic chips, highlighting that the economics can tilt favorably if 3D integration is executed wisely. The next sections

will explore where these trade-offs are being made in practice and how 3D ICs are influencing several major tech sectors.

Applicability of 3D Integrated Circuits

The adoption of 3D IC technology is being driven by its applicability to a range of sectors where improvements in integration density, energy efficiency, and performance is highly valued. Seven major sectors seeing active use and development of 3D integrated circuits are: (a) data centers and high-performance computing, (b) smartphones and mobile devices, and (c) artificial intelligence (AI) hardware, (d) Memory Stacking, (e)Heterogenous SoC, (f) FPGA's and Reconfigurable Computing, (g) Specialized sensors and electronics.

Data Centers and High-Performance Computing (HPC)

Data centers house server and accelerator chips that run our cloud services, internet applications, and large-scale computations. The demand for computing horsepower in data centers is enormous, and energy consumption is a key concern (both for cost and sustainability). 3D ICs have started to play a crucial role in pushing data-center chip performance while keeping power in check. One prominent use of 3D integration in this space is the integration of high-bandwidth memory with processors. Modern GPUs and AI accelerators for data centers often employ 3D-stacked memory (HBM): for example, *NVIDIA's A100* and *H100* data center GPUs and *Google's TPUs* incorporate HBM stacks on the same package as the logic die. Each HBM stack consists of multiple DRAM dies stacked and interconnected by TSVs, providing dramatically

higher memory bandwidth (hundreds of GB/s up to multiple TB/s) directly adjacent to the processor. This reduces the energy per bit versus using external GDDR memory and enables handling the data-intensive workloads of AI training and HPC. The use of HBM is a direct application of 3D IC technology – JEDEC's HBM standard explicitly involves through-silicon vias connecting ~8-12 dies of memory plus an interface die. The performance-per-watt benefits are substantial: one analysis by AMD and SK Hynix showed that an HBM2 module could deliver the same bandwidth as a traditional memory subsystem while consuming far less power, due to shorter interconnects and lower-voltage signaling (exact figures vary, but on the order of 4X improvement in energy efficiency for memory access is often cited). Because of these benefits, virtually all high-end data center accelerators now use 3D-stacked memory – a clear testament to 3D IC value in HPC.

3D integration is also enabling multi-chip modules (MCMs) and chiplet-based server CPUs that break the traditional monolithic processor design. *AMD's EPYC* series of server processors is a prime example: AMD uses a chiplet architecture (branded Infinity Fabric) where each processor package contains multiple smaller CPU core dies (chiplets) and an I/O die, rather than one large die. Early EPYC generations used a 2D MCM (chiplets on a package substrate), but more recently AMD introduced 3D-stacked cache technology. In 2022, AMD launched "Milan-X" EPYC processors featuring 3D V-Cache, where an additional 64 MB SRAM cache die is bonded atop each CPU die, increasing the L3 cache size dramatically for certain workloads. This 3D-stacked cache delivers substantial performance boosts in data center tasks like database queries and

technical computing, by keeping more data on-chip and reducing costly memory fetches. It is achieved through TSV-based hybrid bonding between the cache die and the CPU die. Real-world results showed up to 50–70% performance gains in targeted workloads for Milan-X versus standard Milan chips, with minimal impact on power consumption – an illustration of performance-per-watt improvement via 3D IC technology. AMD's approach underscores how 3D ICs can augment high-performance processors (adding memory closer to cores) in a cost-effective way: the base CPU chiplets are unchanged, and only the extra cache die is added using TSMC's advanced chip stacking process, leveraging existing IP. Similarly, Intel is using 3D integration in its data center products: the upcoming Intel Xeon server processors employ Intel's Foveros 3D packaging to stack a small high-density logic die (containing power management and I/O functions) beneath compute chiplets, effectively integrating what would be two levels of logic in one package. Intel's Ponte Vecchio, a complex HPC/AI chip, uses Foveros to stack compute tiles and EMIB (Embedded Multi-die Interconnect Bridge, a 2.5D interposer) to connect many chiplets and HBM memory into one module – a showcase of heterogeneous integration. These examples highlight data center CPUs and accelerators as a leading frontier for 3D IC adoption, because the economic payoff in performance and efficiency justifies the added complexity. In fact, the AI boom has been a catalyst: analysts note that artificial intelligence workloads are "pushing the limits of high-performance computing and low-power electronics, requiring advances in microelectronics capabilities — especially advanced packaging". The insatiable need for compute in AI training has made technologies like chiplet integration and die stacking increasingly mainstream in data

center hardware. As a result, major companies all have active 3D IC programs for their server and HPC product lines.

Smartphones and Mobile Devices

Mobile devices, particularly smartphones, have arguably the most stringent demands on integration: they require high performance and functionality within a very small power and size envelope. 3D IC technology has found its way into smartphones primarily because of the space savings and memory bandwidth benefits it offers. A smartphone system-on-chip (SoC) must combine application processors, graphics, AI engines, modems, and memory in a tiny package. For years, package-on-package (PoP) assembly has been used in phones – typically the DRAM package is stacked on top of the logic SoC package to shorten connections. This is a form of 3D packaging (albeit with less dense interconnects than TSVs). Today, advanced smartphones are beginning to use true TSV-based 3D stacking. For example, high-end mobile processors integrate stacked DRAM (Wide I/O or similar) to boost memory bandwidth for 4K video and AI applications. *Samsung* has demonstrated an SoC with a 3D-stacked SRAM cache for mobile use, and there are reports that future smartphone APs (application processors) may adopt stacked "chiplet" architectures as Moore's Law slows. One already ubiquitous 3D IC component in phones is the camera sensor: most flagship smartphone camera sensors are stacked CMOS image sensors, where the photodiode array layer is bonded to a separate image signal processor (ISP) layer or memory layer underneath. Sony pioneered this stacked sensor design around 2012 to enable faster readout and on-sensor memory for high-speed photography. This is essentially a 3D IC – two semiconductor

layers (one for pixels, one for logic) connected by fine-pitch bonding – and it has become standard in mobile cameras to achieve high performance in a small module. Thus, even if consumers are not aware of it, 3D integrated circuits are in their pockets, making features like slow-motion video and high-resolution images possible.

Smartphones also benefit greatly from the power efficiency that 3D ICs can provide. Any reduction in power consumption translates to longer battery life or lower heat. By shortening interconnects and by enabling more efficient memory access (through on-chip memory stacking), 3D ICs help mobile designers meet energy budgets. For instance, a 3D-stacked DRAM can operate at lower voltage and higher bandwidth than external memory, meaning the processor can spend less energy waiting on data. Additionally, 3D integration contributes to component miniaturization: freeing up board area or package area for other components like larger batteries or additional sensors. One example at the packaging level is Apple's use of advanced fan-out and likely 2.5D/3D techniques in their M1 Ultra chip: Apple connected two M1 Max chips via a silicon interposer (an approach akin to 2.5D) to act as one larger chip, delivering a multi-die solution in a single package for desktop systems. While that specific example is not a smartphone, it shows how even consumer electronics are embracing multi-chip packaging for performance scaling. We anticipate that as 3D processes become cheaper and more standardized, future mobile SoCs might integrate, for example, a base logic die with various functional chiplet dies (RF, connectivity, AI accelerators) in a single 3D package – effectively an entire phone on a chip. The continued drive for slimmer phones with more features will likely push 3D IC adoption further in this sector.

AI Hardware and Specialized Accelerators

AI hardware spans from giant cloud-based AI training systems to compact edge AI chips in devices like smart cameras or IoT sensors. Across this spectrum, 3D integrated circuits are increasingly important for overcoming the "memory wall" and "communication bottlenecks" that characterize AI workloads. Neural network processing involves multiplying and accumulating vast amounts of data. Performance is often limited by how fast data can be shuttled between memory and compute units, rather than just raw compute speed. 3D ICs offer solutions by bringing memory closer to compute and by enabling massively parallel data paths.

In cutting-edge AI training chips (mostly hosted in data centers, overlapping with the HPC discussion), the use of HBM memory stacks has been a game-changer – providing each GPU/AI processor with hundreds of gigabytes per second of local memory bandwidth. For instance, each *NVIDIA H100 GPU* comes with multiple HBM3 stacks on-package, collectively providing over 3 TB/s of bandwidth, which is critical for large neural network training. This simply would not be feasible without 3D integration; the physical distance and energy cost of communicating with off-package DRAM would make such bandwidth prohibitively power-hungry. Another example in AI accelerators is the development of in-package optical interconnects as chiplets: companies like *Ayar Labs* have created optical I/O chiplets that can be integrated next to processors to communicate between chips with light instead of electrical signals. These optical chiplets, which adhere to the emerging UCIe chiplet interconnect standard, use 3D

integration to achieve extremely high bandwidth (on the order of 8 TB/s in Ayar's chiplet) with low power, addressing the scaling of AI system interconnects beyond what traditional copper links can do. This is a more exotic use of 3D heterogeneous integration tailored for AI scale-out architectures, highlighting how AI needs are driving innovation in 3D IC design beyond just stacking memory.

Even in smaller-scale AI inference chips (like those in smartphones or embedded devices), 3D integration is making inroads. One approach under exploration is logic-on-memory stacking for neural network accelerators, where compute layers are stacked directly atop dense memory arrays (e.g., RRAM or SRAM) to form 3D "neuronal" tiles, minimizing data movement for matrix operations. Research prototypes have shown orders-of-magnitude improvement in energy efficiency by using monolithic 3D integration to intermix logic and memory at a fine grain (sometimes called "neuromorphic 3D ICs"). While these are still in the R&D phase, they indicate the path forward for AI hardware: integrating computing and memory elements vertically to mimic the efficiency of biological systems and break the memory bottleneck.

In summary, from massive cloud AI engines to tiny edge AI chips, 3D ICs are becoming a key enabling technology. They allow AI accelerators to scale in performance (handling larger vhardware players like NVIDIA, Google, AMD, and Cerebras are all leveraging advanced packaging. The extreme performance requirements of AI make it one of the most important drivers of 3D IC technology today.

Memory Stacking

As discussed in the historical overview, memory was an early adopter of 3D technology. NAND flash memory manufacturers stack dozens of dies to achieve high capacity in a small package (e.g., 8, 16 or even 32 flash dies connected by wire bond or TSV in one package for SSDs and memory cards). This is distinct from 3D NAND technology (where memory cells themselves are built vertically), but often both approaches are combined – for instance, Samsung's 512 GB flash chip (2017) stacked eight dies, each of which was a 64-layer 3D NAND device, for an effective "8 x 64-layer" flash stack. For volatile memory, DDR DRAM modules using TSVs began appearing around 2014–2015 in the form of stacked DRAM for servers and high-performance systems. A key breakthrough was the standardization of High Bandwidth Memory (HBM), which places 4–8 DRAM dies on top of a logic interface die. The logic dies handles I/O and is connected to the stack of DRAM dies by thousands of TSVs carrying data, address, and control lines. The entire stack is then typically mounted on an interposer alongside a host processor or GPU. HBM enables very wide data interfaces (e.g. 1024-bit) running at moderate speed, yielding enormous aggregate bandwidth. For example, the first generation HBM provided ~128 GB/s per stack, and later HBM2/2E stacks provide 256 GB/s or more, which has been crucial for GPU performance in graphics and AI workloads. 3D-stacked memory like HBM or the earlier Hybrid Memory Cube (HMC) are now common in applications requiring fast memory access, from high-end graphics cards to network routers.

Heterogeneous System-on-Chip

3D ICs open the possibility of integrating different types of functionalities that would traditionally require separate chips. One major application is logic-memory integration. By putting memory on top of logic, one can achieve bandwidth and latency improvements impossible in 2D. A prominent example is processor with stacked L3 cache – AMD's 3D V-Cache technology stacks a 64 MB SRAM die atop a Ryzen CPU die, increasing the cache size without enlarging the CPU chip and boosting performance for cache-sensitive applications. Another example is advanced mobile SoCs that have stacked DRAM. Image sensors are also a success story of heterogeneous 3D integration: modern CMOS image sensors frequently use a stacked architecture where the pixel array is on one die and the image signal processor is on a second die beneath it. The pixel die can be optimized in a specialized process and thinned to enable back-illumination, while the logic die can be a standard CMOS process. TSVs or micro-bump arrays connect the pixel outputs to the logic die below. This 3D structure greatly improves camera performance in smartphones and digital cameras by enabling more complex processing right under the pixel array and reducing interconnect length.

FPGAs and Reconfigurable Computing

FPGAs benefit from 3D integration primarily through 2.5D implementations at present (like Xilinx's multi-die FPGAs). Looking forward, one could imagine stacking an FPGA fabric die with a mass of configuration memory on a separate die – since FPGAs are interconnect-dominated, a 3D architecture might drastically reduce routing delays by having multiple active routing layers. For commercial products, Xilinx's 2.5D approach

using an interposer (first deployed in Virtex-7 2000T and later devices) demonstrated the effectiveness of connecting four large FPGA dies on a silicon. This is sometimes termed "pseudo-3D" but it addresses similar goals: it allowed Xilinx to scale FPGA capacity using multiple smaller chips rather than one huge die, improving yield and cost. Future FPGA generations may further use 3D by stacking flash or FeRAM for configuration memory or mixing and matching process nodes (an FPGA fabric die made in a logic-optimized node with a transceiver/analog die in a RF-optimized node, etc., all in one package).

Specialized Sensors and Electronics

3D ICs have been used in niche applications like high-energy physics instrumentation, where sensors, analog front ends, and digital processing are stacked to achieve ultra-compact detectors. For example, a "3D silicon pixel detector" can have a sensor layer, an analog readout layer, and a digital layer, enabling real-time processing at the pixel level – an architecture investigated for particle trackers and imaging arrays. Similarly, memory cubes were applied in some networking and computing systems to provide deep 3D memory modules for data centers.

3D ICs in National Security and Semiconductor Sovereignty

Given the strategic importance of semiconductors to economic and military power, many governments have recognized advanced chip packaging and 3D integration as critical technologies for national security and supply chain resilience. In recent years, initiatives in the United States, Europe, and Asia (notably China, Taiwan, and South

Korea) have poured resources into ensuring leadership or self-sufficiency in advanced packaging, including 3D IC capabilities. The rationale is that even if a country leads in chip fabrication , leadership in packaging and integration is equally vital to assemble those chips into cutting-edge systems. Below we discuss how 3D ICs feature in these national and international strategies, highlighting the U.S. CHIPS Act and other programs in China.

United States – CHIPS Act and DARPA Initiatives

The United States has launched major funding efforts to revitalize domestic semiconductor manufacturing with an emphasis not only on fab processes but also on packaging. The 2022 CHIPS and Science Act authorized significant funding (over $50 billion) to boost the U.S. semiconductor industry. Out of this, approximately $3 billion is earmarked for a *National Advanced Packaging Manufacturing Program (NAPMP)* aimed at making the U.S. a leader in 2.5D and 3D packaging technology. In late 2023, the U.S. Commerce Department released a vision for this program, stating that within a decade, they envision America will manufacture and package the world's most sophisticated chips, onshoring a high-volume advanced packaging industry that is self-sustaining. This statement highlights that advanced packaging (which includes 3D IC assembly) is viewed as a national priority, critical to a "thriving semiconductor ecosystem". Concretely, CHIPS Act funding is being used to establish new R&D centers and pilot lines for 3D integration. In 2024, the Department of Commerce launched a competition for $1.6 billion in R&D funding under NAPMP to spur innovation across five areas of advanced packaging, with plans to award around $150 million to several research centers in each

area. At the same time, the Department of Defense (DoD) is investing in this space through its "Microelectronics Commons" program, funding regional hubs focusing on bridging the gap from lab research to fab prototyping for technologies including heterogeneous 3D integration. In 2023, DoD announced $238 million for 8 regional innovation hubs under this program, targeting areas like 5G/6G, AI hardware, and "secure edge" computing – all of which involve advanced integrated microsystems.

Perhaps the most direct indicator of 3D ICs' strategic value is DARPA's recent programs. In mid-2024, DARPA awarded an $840 million project to a consortium led by the University of Texas (Texas Institute for Electronics) to develop next-generation 3D heterogeneous integration technologies for the U.S. military. This enormous contract, under *DARPA's Electronics Resurgence Initiative*, aims to create a manufacturing center for stacking layers of silicon dies on top of each other to achieve leaps in performance and integration density for defense applications. The focus on 3D Heterogenous Integration suggests the DoD sees vertically integrated chips as essential for future defense systems – from compact, high-performance sensors and AI processors in drones, to secure computing modules that can't be easily replicated by adversaries. Notably, DARPA had earlier funded the "CHIPS" program to develop chiplet standards, and the "3DSoC" program for monolithic 3D SoCs, indicating a long-term interest in breaking the 2D mold. The fact that U.S. defense research is dedicating such resources to 3D chips underscores their strategic importance. By investing now, the U.S. hopes to avoid being dependent on foreign packaging providers and to secure access to the best integration tech for its own use. There is also an offensive angle: if the U.S. can leap ahead in 3D IC

capability, its semiconductor industry could outpace others even if raw transistor advances slow down.

China – Advanced Packaging to Close the Gap

China's situation in semiconductors is unique due to export controls on leading-edge fab equipment and its own national goal of self-reliance. Unable to manufacture 5 nm and below chips at scale domestically, China is investing heavily in advanced packaging and chiplet technologies to maximize the performance of the nodes it can produce (e.g., 14 nm, 7 nm with difficulty). The logic is that if Chinese companies can stack and interconnect chips skillfully, they might achieve performance comparable to a single 3 nm chip by using multiple 14 nm chips, for instance. We see evidence of this strategy: Chinese OSATs like JCET and Huatian are pouring resources into 2.5D interposers, fan-out-on-substrate, and 3D stacking capabilities. In 2021, TSMC's Chinese competitor SMIC announced a plan to develop "TSV 3D packaging" techniques. Huawei's HiSilicon, which lost access to TSMC's 7 nm after sanctions, reportedly turned to chiplet designs to link multiple mid-range chips together to approximate a high-end chip for applications like servers and base stations. Moreover, China's national semiconductor fund has invested in advanced packaging houses, recognizing that packaging could be a comparative advantage if cutting-edge lithography is a bottleneck. Chinese research institutions are also active in 3D IC R&D; for example, Tsinghua University and others publish on 3D architectures and have built prototypes like a 3D AI chip with on-chip optical interconnects.

On the national security front, China likely views mastery of 3D ICs as essential for its military electronics to be on par with U.S. systems. Secure and radiation-hardened electronics, high-performance chips for communications and AI in defense – all these could benefit from 3D integration. However, one challenge China faces is that some advanced packaging equipment and EDA tools are also subject to export restrictions. This underscores that global powers see 3D IC tech as a strategic asset worth controlling. Nonetheless, China is pushing forward through domestic innovation and by leveraging open standards (like the UCIe consortium, which Alibaba joined as a board member). We can expect China to use 3D ICs to leapfrog in performance where they might lag in single-die technology – for instance, using a chiplet approach to build an AI training engine from multiple chips integrated on an advanced packaging substrate, rather than one monolithic 5 nm GPU which they cannot produce internally. The outcomes of these efforts remain to be seen, but clearly 3D integration has become a key battleground in the race for semiconductor independence.

In summary, at the geopolitical level, 3D IC technology is considered a strategic frontier. The U.S. is heavily investing to regain leadership and secure its supply and China is focusing on packaging to mitigate its weaknesses in fabrication. Taiwan and South Korea, for completeness, are also in the mix: TSMC (Taiwan) and Samsung (Korea) are not standing still – they lead many 3D packaging innovations commercially (TSMC's CoWoS and SoIC technologies are enabling customers like Apple, AMD, and Nvidia, and Samsung's X-Cube is offering 3D stacking to its foundry clients). These companies often collaborate with their governments (e.g., Taiwan has a national advanced packaging center, and Samsung receives government support for packaging

R&D). We are witnessing a “more-than-Moore” space race, where success in 3D ICs could translate to economic gains and technological advantage in critical sectors like artificial intelligence, communications, and defense.

Investment and R&D Trends in 3D IC Technologies

The growing importance of 3D integrated circuits is mirrored by significant trends in industry investments, venture capital funding, and collaborative R&D efforts targeting these technologies. In this section, we outline how companies and investors are driving 3D IC advancement and the formation of an ecosystem to support it. Key trends include substantial capital expenditures by semiconductor firms on advanced packaging facilities; increasing venture funding for startups focusing on chiplet interconnects, 3D packaging materials, and related innovations; enhanced R&D collaboration through consortia and standards bodies (to solve common challenges in 3D design and manufacturing); and the expansion of EDA and manufacturing tool support for 3D integration.

Corporate Investments and Capacity Expansion

Leading semiconductor manufacturers and assembly houses have been investing heavily to build up 3D IC production capabilities. One notable example is Taiwan’s *ASE* (Advanced Semiconductor Engineering) – the world’s largest OSAT – which, according to industry reports, is roughly doubling its advanced packaging capacity in Silicon Valley to serve clients like *AMD, Apple, and Nvidia* with 3D integration and advanced chip packaging services. This expansion, aided by a $1.6B CHIPS Act subsidy, shows that demand for 3D packaging is surging to the point where existing capacity was insufficient.

Similarly, TSMC has been ramping up its "3DFabric" portfolio – building new facilities dedicated to these technologies, both in Taiwan and abroad (for instance, the company's Arizona fab project is rumored to include an advanced packaging plant to assemble chiplets from the fab). Intel, as both an IDM and a nascent foundry, has made *Foveros* 3D packaging a centerpiece of its roadmap, investing in sites in Oregon, Arizona, and Malaysia to handle volume manufacturing of *Foveros*-stacked chips. Intel even opened its advanced packaging facility to third-party customers, betting that offering cutting-edge 3D integration will be a competitive differentiator for its foundry services. On the memory side, companies like Samsung and SK Hynix have invested in HBM production lines – essentially advanced 3D stacking of DRAM dies – to meet the explosive demand from AI markets. These moves indicate that big semiconductor firms see return on investment in 3D IC technology and are allocating billion-dollar budgets to ensure they have the needed infrastructure. This observation is in line with the flurry of current investments: whoever builds the capacity and know-how first will attract the next wave of chip projects that need 3D integration.

Venture Funding and Startups

While much of the 3D IC revolution is driven by established giants, there is also a vibrant startup scene tackling specific pieces of the 3D puzzle. Venture capital has taken interest in areas like chiplet interconnect standards, novel bonding techniques, and optical or wireless die-to-die communication – all enabling technologies for easier 3D integration. For example, *Eliyan*, a startup focused on chiplet interconnect IP (the physical interface that allows chiplets to communicate within a package), raised

substantial funding (tens of millions of dollars) to develop its high-speed "NuLink" interconnect, which aims to simplify connecting chiplets without an expensive interposer. Another startup, Ayar Labs, which I mentioned earlier, has raised over $100M in venture funding to develop optical I/O chiplets that can be integrated via the UCIe standard for AI and HPC applications. Startups like zGlue and Chipletz have explored marketplaces for chiplets and Lego-like integration of small dies for IoT devices, receiving seed funding. There are also materials and equipment startups – for instance, firms developing new dielectric adhesives or thermal interface materials suitable for 3D stacking, or companies working on wafer-to-wafer bonding equipment that can achieve better alignment and throughput. The presence of these startups is significant: they often drive innovation that larger companies later adopt or acquire. It also shows that investors believe there is a growing market for tools and IP that facilitate heterogeneous integration. As more case studies prove the value of 3D ICs, we can expect venture funding to further accelerate, especially in standardizing chiplet ecosystems. The formation of a robust ecosystem could create entirely new business models and markets, akin to how the IP core business flourished during the SoC era.

Collaborative R&D and Consortia

Another trend is the rise of collaborative efforts to develop 3D IC design methodologies and standards. In March 2022, a group of industry leaders announced the *Universal Chiplet Interconnect Express* (UCIe) consortium, an open industry standard for die-to-die interconnect at the package level. The founding board members of UCIe include *Intel, AMD, Arm, TSMC, Samsung, ASE, Qualcomm, Microsoft, Meta,* and later

Google and *Alibaba* joined. The UCIe standard defines a common communication protocol and physical interface so that chiplets from different vendors can interoperate if placed in the same package. This is a crucial step toward a marketplace of mix-and-match chiplets. The consortium's rapid growth (over 60 companies by late 2022) and the release of version 1.1 and 2.0 of the specification show the momentum behind standardizing 3D integration. Similarly, the *Open Compute Project (*OCP*)* established the *Chiplet Design Exchange (*CDX*)* in 2021, bringing together EDA companies, fabs, OSATs, and system vendors to define data models and design kits for chiplets. The goal is to make designing a 3D IC as close as possible to the familiar flow of designing a PCB with multiple components – a high-level assembly approach, but for chips. These collaborations are important because they address one of the barriers to wider adoption: the lack of standardized design flows and business practices for integrating IP from different sources in one package. If successful, they will reduce the entry barrier and cost for smaller companies to adopt 3D ICs (since they could license a chiplet for a function instead of designing it, and know it will work with others via UCIe, for example). The industry consensus behind UCIe also means that even fierce competitors agree on the need for a common interconnect – a sign that the chiplet model is expected to be ubiquitous and that no single company can unilaterally dominate it without standards.

On the research side, universities and government labs are partnering with industry under various consortia to solve technical challenges of 3D ICs, such as thermal management (e.g., DARPA's 3DSoC program involved academic teams to devise low-temperature processes for monolithic 3D), or testing (the IEEE hosted special sessions on

3D IC test, often concluding that design-for-test collaboration is needed between chip vendors and equipment makers). One example is the MIT Lincoln Laboratory and DARPA collaboration that demonstrated a 3D chip with stacked logic and memory for intelligence applications, sharing lessons on yield and security of 3D chips. The SRC (Semiconductor Research Corporation) also has a program on Heterogeneous Integration that funds university research in 3D packaging, co-sponsored by companies and the U.S. government. All these R&D efforts feed into the ecosystem, ensuring a pipeline of new ideas and skilled engineers for the 3D IC domain.

EDA and Toolchain Development

To support the rise of 3D ICs, electronic design automation vendors (Cadence, Synopsys, Siemens EDA, Ansys, etc.) have been actively developing new tool capabilities. These include tools for chiplet-based floorplanning, thermal and stress simulation for stacked dies, and 3D-aware timing analysis. Siemens and Ansys, for example, have emphasized solutions for thermal-electrical co-simulation given the power density issues of 3D ICs. Synopsys and Cadence have introduced “2.5D/3D IC design suites” that integrate what used to be separate domains: chip design and package design, into one environment. This is crucial for system technology co-optimization – essentially treating the multi-die package as an extended chip that needs holistic optimization. The availability of better design tools lowers the risk and effort for companies to try 3D integration, thereby encouraging more adoption. We have already seen startups leveraging this; for instance, some fabless companies that don’t have internal packaging

expertise can now use EDA tools to virtually prototype a 3D IC and then outsource the assembly to OSATs with confidence that it will meet specifications.

In summary, the landscape of investment and R&D in 3D ICs is vibrant and expanding. Big industry players are betting on 3D integration as the next driver of growth, startups and venture capital are injecting innovation and helping create the ecosystem of supporting technologies, and collaborative initiatives are aligning the industry on common standards and knowledge sharing. A concrete sign of the times is that trade conferences like IEEE's Electronic Components and Technology Conference or the 3DInCites Forum have seen record participation, with topics like chiplet standards, hybrid bonding, and 3D system design dominating the agenda. The trend is clear: 3D ICs are transitioning from niche technology to mainstream methodology. This has put some urgency into the rest of the industry to catch up, fueling even more investment – no one wants to be left behind if 3D ICs define the next decade of electronics. The confluence of funding, talent, and strategic necessity is creating a virtuous cycle propelling 3D IC innovation forward.

Conclusion

This thesis has presented a comprehensive study of 3D integrated circuits, examining their technical foundations, fabrication processes, economic implications, and strategic significance for the semiconductor industry. It began by establishing the technical underpinnings of 3D IC technology, describing how techniques like through-silicon vias and wafer bonding make vertical chip stacking possible. The thesis surveyed

state-of-the-art manufacturing processes – from via-first approaches in wafer fabrication to die-to-wafer bonding techniques – and discussed how these processes have evolved to enable high-density 3D integration. In addition to the engineering aspects, we analyzed the economic and business factors influencing 3D IC adoption, including cost trade-offs, yield considerations, and supply chain dynamics Through this multifaceted investigation, the thesis has elucidated the profound potential of 3D ICs as well as the practical challenges that must be addressed to realize that potential.

One of the main contributions of this work is an integrated perspective that combines both technical and business insights. Technologically, 3D ICs represent a paradigm shift in chip design – offering a path to continue increasing device density and performance beyond what conventional planar scaling can achieve. We have shown that by stacking chips, significant improvements in performance and energy efficiency are attainable: for example, memory-on-logic stacking can provide massive bandwidth boosts while cutting power consumption per data transfer. Such advances directly tackle the bottlenecks that are emerging as Moore's Law slows and Dennard scaling has ended. Instead of relying solely on smaller transistors, 3D integration brings a new dimension to maintain the trajectory of computational capability growth. This thesis underscored how 3D ICs can alleviate on-chip interconnect delays, reduce signal travel distances, and enable new architectures – effectively offering an avenue to extend Moore's Law by other means. At the same time, our economic and strategic analysis reinforces that the transition to 3D ICs is not just a technical upgrade but a strategic imperative for the industry. As scaling costs rise and returns diminish, companies must weigh the

investment in 3D integration against the risk of stagnation in product performance. Here, the thesis argued that the transformative potential of 3D ICs – in meeting emerging computational demands like artificial intelligence and big data, and in enabling continued miniaturization for mobile and IoT devices – makes a compelling case for continued R&D and capital expenditure in this area. The economic analysis showed that while upfront costs are high, the long-term benefits in product differentiation and capability can justify the investment, especially as ecosystem support for 3D ICs grows and economies of scale improve.

Crucially, this work has highlighted that the success of 3D IC technology will rely on holistically integrating both technical innovation and business strategy. Technical solutions to thermal issues, yield improvement, and design complexity must go together with industry collaboration on standards, supply chain development, and cost reduction. In other words, solving the engineering problems alone is not enough; the ecosystem — from EDA tool providers to foundries, OSATs, and device makers — needs to align to support 3D IC adoption on a broad scale. This thesis has illustrated how such integration of perspectives can inform better decision-making. For instance, understanding the economic value of a performance gain can guide where to focus engineering effort. Conversely, knowing the technical constraints can prevent over-investing in a business direction that current science cannot yet deliver. By reflecting on both the technical feasibility and the commercial viability of 3D ICs, the thesis makes a strong argument that continued investment and innovation in this field are not only warranted but necessary. The challenges we identified – from thermal management to high costs – are

significant but are surmountable through coordinated innovation, as evidenced by the rapid progress in recent years. Each year, new milestones are being reached: higher stacking counts, new memory-on-logic products, better EDA support, and growing industry adoption. These trends affirm that 3D ICs are transitioning from experimental technology into a mainstream solution for computing needs.

In conclusion, 3D integrated circuits stand at the forefront of a new era for the semiconductor industry. They represent a paradigm shift in how we think about scaling: moving "upwards" instead of outwards and integrating functionalities in ways previously unimaginable on a flat silicon die. The research and analyses presented have shown that 3D ICs hold tremendous promise in addressing the grand challenges facing electronics today, from the plateauing of Moore's Law to the exploding demand for computing performance and memory bandwidth in modern applications. The ability to pack more transistors and memory into a smaller volume, to cut down communication latencies, and to create heterogeneous systems by stacking different technologies, all points to 3D ICs as a transformative engine of innovation in the coming decades. There is a clear rationale for the industry to continue pushing this technology forward: the benefits in energy-efficient high-performance computing, miniaturization of devices, and even new capabilities like built-in hardware security, are too important to ignore. As we look to the future, it is evident that the role of 3D ICs will only grow more critical. We can foresee a time when vertically integrated chips become commonplace in servers powering the cloud, in AI supercomputers driving scientific discovery, and in the ubiquitous smart devices of daily life. In shaping that future, the insights from both engineering and

economics must guide the way. The findings of this thesis reinforce that with sustained effort in research, development, and collaboration across the semiconductor ecosystem, 3D integrated circuits will play a pivotal role in shaping the future of computing, ensuring that the momentum of progress continues even as traditional scaling slows. In the grand tapestry of the semiconductor industry's evolution, 3D ICs are poised to be a key thread, weaving together advances that keep us on the cutting edge of technology for years to come.

Glossary of Terms and Abbreviations

AI – Artificial Intelligence: Algorithms and systems that perform tasks typically requiring human intelligence.

ALU – Arithmetic Logic Unit: A digital circuit used to perform arithmetic and logic operations in processors.

AMD – Advanced Micro Devices: A leading semiconductor company known for CPUs and GPUs.

Ansys – A company that develops multiphysics simulation software used in chip design and verification.

APs – Application Processors: Processors designed to support applications running on devices like smartphones or tablets.

Apple M1 Max/Ultra Chip – High-performance chips from Apple designed with unified memory architecture and 3D integration.

ASE – Advanced Semiconductor Engineering: A leading provider of semiconductor manufacturing and testing services.

BIST – Built-In Self-Test: A design technique that allows a chip to test itself for faults.

Cadence – A company offering EDA (Electronic Design Automation) tools for semiconductor design.

CDX – Chiplet Design Exchange: An initiative aimed at standardizing chiplet interfaces and integration.

CHIPS Act – Creating Helpful Incentives to Produce Semiconductors Act: U.S. legislation to boost domestic semiconductor manufacturing.

Chiplet – A smaller functional block of an integrated circuit designed to be combined with others in a larger package.

CMOS – Complementary Metal-Oxide-Semiconductor: A technology for constructing integrated circuits.

CPU – Central Processing Unit: The main processor in a computer that performs most of the processing tasks.

DARPA – Defense Advanced Research Projects Agency: A U.S. agency supporting high-tech research including semiconductors.

DDR – Double Data Rate: A type of DRAM that transfers data on both the rising and falling edges of the clock signal.

DDR3 SDRAM – Third-generation DDR Synchronous DRAM, offering improved performance over earlier generations.

Dennard Scaling – A principle stating that power density remains constant as transistors shrink, now largely invalid at nanoscales.

DRAM – Dynamic Random-Access Memory: A type of volatile memory used in computing devices.

EDA Tool – Electronic Design Automation Tool: Software used to design and verify ICs and PCBs.

Elpida – A former Japanese DRAM manufacturer acquired by Micron.

EMIB – Embedded Multi-die Interconnect Bridge: Intel's 2.5D packaging technology for connecting dies.

Fabs – Fabrication Facilities: Plants where semiconductor devices are manufactured.

Fan-out-on-substrate – A packaging method enabling high-density interconnects by redistributing I/O pads.

Floor planning – The process of arranging functional blocks of an IC design to optimize performance and manufacturability.

FLOP – Floating Point Operation: A measure of computer performance, especially in scientific calculations.

Foundries – Companies that manufacture chips designed by other firms (e.g., TSMC, GlobalFoundries).

FPGA – Field Programmable Gate Array: A reconfigurable IC that can be programmed after manufacturing.

GDDR – Graphics Double Data Rate: A type of DRAM used in graphics cards.

GlobalFoundries – A semiconductor foundry that manufactures integrated circuits for various clients.

Google TPU – Google's Tensor Processing Unit: A chip optimized for machine learning workloads.

GPU – Graphics Processing Unit: A processor optimized for parallel processing, essential in graphics and AI.

HBM – High Bandwidth Memory: A high-speed memory standard used in GPUs and AI accelerators.

HBM3 – The third generation of High Bandwidth Memory with enhanced performance and bandwidth.

Heterogeneous Systems – Systems integrating different types of processors (e.g., CPU, GPU, FPGA) or technologies (e.g., memory, logic).

HiSilicon – A semiconductor company and subsidiary of Huawei Technologies.

HMC – Hybrid Memory Cube: A 3D-stacked DRAM product with a logic layer for memory control.

HPC – High-Performance Computing: The use of supercomputers and parallel processing for advanced computations.

Huawei – A multinational tech company that develops consumer electronics and semiconductor products.

Huatian – A provider of semiconductor packaging and testing services.

I/O – Input/Output: Interfaces for communication between a device and other components or systems.

IEEE ECTC – IEEE Electronic Components and Technology Conference: A leading conference on electronics packaging and systems.

IDM – Integrated Device Manufacturer: A company that designs, manufactures, and sells ICs (e.g., Intel).

Intel – A major semiconductor company known for its CPUs and 3D integration tech like Foveros.

Intel Foveros – Intel's 3D packaging technology using die stacking and through-silicon vias.

IR Drops – Voltage drops in power delivery networks due to resistance.

IOT – Internet of Things: A network of physical objects embedded with sensors and connectivity.

JCET – Jiangsu Changjiang Electronics Technology: A Chinese OSAT company.

JEDEC – Joint Electron Device Engineering Council: The semiconductor industry's standards body.

JTAG – Joint Test Action Group: A standard for testing and debugging ICs.

Ldi/dt – A measure of the rate of change of current over time, relevant in power integrity analysis.

LSI Chip – Large-Scale Integration Chip: ICs that contain thousands of transistors.

Memory-on-logic products – Devices that stack memory directly above logic to reduce latency and power consumption.

Meta (Facebook) – A tech company investing in AI and custom semiconductor development.

Micron – A U.S. semiconductor company specializing in memory products like DRAM and NAND.

Microsoft – A major technology company investing in custom silicon for AI and cloud.

MOS IC – Metal-Oxide-Semiconductor Integrated Circuit: A fundamental IC structure.

MCM – Multi-Chip Module: A packaging solution where multiple ICs are placed into a single package.

Nascent Foundry – A new or emerging semiconductor foundry.

NVIDIA – A semiconductor company known for GPUs and AI accelerators.

NVIDIA A100/H100 – High-end GPUs from NVIDIA designed for AI and HPC workloads.

Neural Network Accelerators – Specialized processors optimized for machine learning inference/training.

Neuronal / Neuromorphic 3-D IC – ICs mimicking the structure of the human brain for efficient AI processing.

NMOS – N-type Metal-Oxide-Semiconductor: A type of transistor.

NoC – Network-on-Chip: An interconnect scheme used in multi-core processors.

Open Compute Project (OCP) – A collaborative community creating open-source designs for data centers.

OSATs – Outsourced Semiconductor Assembly and Test providers: Companies that offer post-fabrication services.

Package-on-Package (PoP) – A packaging technique where multiple IC packages are stacked.

PCB – Printed Circuit Board: A platform for mechanically supporting and electrically connecting components.

Qualcomm – A company known for mobile processors and wireless communications chips.

RRAM – Resistive RAM: A type of non-volatile memory using resistance change to store data.

RF – Radio Frequency: High-frequency signals used in wireless communication.

Samsung – A major semiconductor company and leader in memory and logic chip manufacturing.

Siemens ADA – Siemens' Advanced Design Automation tools for IC design.

Silicon die – A small block of semiconducting material containing an integrated circuit.

SK Hynix – A South Korean memory semiconductor supplier.

SMIC – Semiconductor Manufacturing International Corporation: China's largest semiconductor foundry.

SoC – System on Chip: An IC that integrates all components of a computer or electronic system.

SRAM – Static Random-Access Memory: Faster but more expensive and power-consuming than DRAM.

Stacking counts – Refers to the number of vertically integrated layers in a 3D IC.

Synopsys – A leading EDA company offering tools for IC design and verification.

TB – Terabyte: A unit of digital data equal to 1,024 gigabytes.

Tezzaron Semiconductor – A company specializing in 3D IC technology and memory integration.

Texas Instruments – A U.S. semiconductor company with legacy in analog and digital ICs.

TSMC – Taiwan Semiconductor Manufacturing Company: The world's largest pure-play foundry.

TSMC SoIC – TSMC's 3D integration platform using chip stacking and wafer bonding.

UCIe – Universal Chiplet Interconnect Express: An open standard for die-to-die interconnects.

Wafer bonding – A technique for joining two semiconductor wafers for 3D integration.

References

“Advanced 2.5D/3D Packaging Roadmap.” *SEMI*, 31 July 2023. PDF.

Advanced Micro Devices. *3rd Gen AMD EPYC Processors with AMD 3D V-Cache Technology.* Product brief, 2022, ir.amd.com/news-events/press-releases/detail/1055/3rd-gen-amd-epyc-processors-with-amd-3d-v-cache-technology. Accessed 9 Mar. 2025.

“Advanced Packaging.” *Semiconductor Engineering*, n.d., semiengineering.com/knowledge-centers/advanced-packaging/. Accessed 13 Nov. 2024.

“Advanced Packaging: 3D TSV Technologies—Via First or Via Last?” Yole Développement, 2010, i-micronews.com. Archived 17 May 2014. Accessed 7 Feb. 2025.

AnandTech. "Intel Unveils a Groundbreaking Way to Make 3D Chips." *Engadget*, 8 Aug. 2019, engadget.com/intel-groundbreaking-3d-chips-method. Accessed 26 Feb. 2025.

Shilov, Anton. "Samsung Starts Production of 512 GB UFS NAND Flash Memory: 64-Layer V-NAND, 860 MB/s Reads." *AnandTech*, 5 Dec. 2017, anandtech.com/show/12133/. Accessed 27 Dec. 2024.

Tallis, Billy. "Samsung Shares SSD Roadmap for QLC NAND and 96-Layer 3D NAND." *AnandTech*, 17 Oct. 2018, anandtech.com/show/13419/. Accessed 11 Jan. 2025.

Bailey, Brian. "3D-IC for the Masses." *Semiconductor Engineering*, 13 Mar. 2025, semiengineering.com/3d-ic-for-the-masses/. Accessed 3 Dec. 2025.

Banerjee, Kaustav, et al. "3-D ICs: A Novel Chip Design for Improving Deep-Submicrometer Interconnect Performance and System-on-Chip Integration." *Proceedings of the IEEE*, vol. 89, no. 5, 2001, pp. 602-33.

Black, B., et al. "3D Processing Technology and Its Impact on iA32 Microprocessors." *Proceedings of the International Conference on Computer Design*, 2004, pp. 316-18.

Bollmann, D., et al. "Three-Dimensional Metallization for Vertically Integrated Circuits." *Microelectronic Engineering*, vols. 37–38, Nov. 1997, pp. 39-47, doi:10.1016/S0167-9317(97)00092-0.

Borkar, Shekhar. "3D Integration for Energy-Efficient System Design." *Proceedings of the Design Automation Conference*, 2011, pp. 214-19.

Chen, D. Y., et al. "Enabling 3D-IC Foundry Technologies for 28 nm Node and Beyond: Through-Silicon-Via Integration with High-Throughput Die-to-Wafer Stacking." *IEEE International Electron Devices Meeting (IEDM)*, 2009, pp. 1-4, doi:10.1109/IEDM.2009.5424350.

Clark, Colin. "3D-IC for the Masses." *Semiconductor Engineering*, 13 Mar. 2023, semiengineering.com/3d-ic-for-the-masses/. Accessed 27 Dec. 2024.

Cole, Bernard. "Terrazon Applies 3D Stacking Technology to 8051 MCU Core." *EE Times*, 22 May 2005, eetimes.com/terrazon-3d-stacking-8051. Accessed 21 Oct. 2024.

Connatser, Matthew. "DARPA Slaps Down Credit Card for 3D Military Chiplets—$840 M Ought to Be Enough?" *The Register*, 18 July 2024, theregister.com/2024/07/18/darpa_awards_840m_to_utaustin/. Accessed 16 Jan. 2025.

Dally, William J. "Future Directions for On-Chip Interconnection Networks." Computer Systems Laboratory, Stanford U., 2006. Archived 12 June 2010. Accessed 22 Mar. 2025.

Dong, Xiangyu, and Yuan Xie. "System-Level Cost Analysis and Design Exploration for 3D ICs." *Asia and South Pacific Design Automation Conference*, 2009. xie.ca.pitt.edu. Archived 24 Apr. 2010. Accessed 13 Feb. 2025.

Dong, Xiangyu, et al. "An Optimized 3D-Stacked Memory Architecture by Exploiting Excessive, High-Density TSV Bandwidth." *Proceedings of the 16th International Symposium on High-Performance Computer Architecture*, IEEE, Jan. 2010, pp. 429-40.

"EDA's Big Three Unready for 3D Chip Packaging." *EE Times Asia*, 25 Oct. 2007. Archived 18 July 2008. Accessed 13 Feb. 2025.

Frank, David J. "Three-Dimensional Integrated Circuits." *IBM Journal of Research and Development*, 2006.

Fukushima, Takashi, and Mitsuhiro Koyanagi. "Thermal Issues of 3D ICs." *SEMATECH 3D IC Symposium*, 2007.

Garg, Sudeep, and Diana Marculescu. "3D-GCP: An Analytical Model for the Impact of Process Variations on the Critical Path Delay Distribution of 3D ICs." *Proceedings of the International Symposium on Quality Electronic Design (ISQED)*, 2009, pp. 147-55.

Laird, Jeremy. "Google Says Per-Transistor Chip Costs Stopped Getting Cheaper 10 Years Ago—But How Is That Right?" *PC Gamer*, 2 Feb. 2024, pcgamer.com/google-says-per-transistor-chip-costs-stopped-getting-cheaper-10-years-ago-but-how. Accessed 6 Nov. 2024.

Hu, Chia-Chi, et al. "3D Integrated Circuit Cooling with Microfluidics." *Micromachines*, vol. 9, no. 6, 2018, article 287, doi:10.3390/mi9060287.

Hynix. "Hynix Surprises NAND Chip Industry." *The Korea Times*, 5 Sept. 2007. Accessed 11 Apr. 2025.

Hynix, SK. "HBM (High Bandwidth Memory) DRAM—Overview and Introduction." Press release, 2015.

IEEE Standards Association. *IEEE Std 1838-2019: Design-for-Test for 3D Stacked ICs.* IEEE, 2019.

Imec. "A Design-for-Test Standard for Test Access in 3D Integrated Circuits." *Imec Magazine*, Mar. 2020, imec-int.com/imec-magazine/2020-march/dft-3d-ics. Accessed 27 Mar. 2025.

"International Technology Roadmap for Semiconductors: 2011 Edition." ITRS, 2011, web.archive.org/web/20150924100353/http://www.semi.org/cms/groups/public/documents/web_content/ctr_033139.pdf. Accessed 21 Oct. 2025.

Jacob, Bruce, et al. “Predicting the Performance of a 3D Processor-Memory Chip Stack.” *IEEE Design & Test of Computers*, vol. 22, no. 6, 2005, pp. 540-47.

James, Dick. “3D ICs in the Real World.” *Proceedings of the 25th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC 2014)*, IEEE, 2014, pp. 113-19, doi:10.1109/ASMC.2014.6846988.

Johnson, R. Colin. “3-D Chip Stacks Standardized.” *EE Times*, 10 July 2008. Archived 30 Sept. 2012. Accessed 13 Dec. 2024.

Kada, Morihiro. “Research and Development History of Three-Dimensional Integration Technology.” *Three-Dimensional Integration of Semiconductors: Processing, Materials, and Applications*, Springer, 2015, pp. 8-13.

Kawamura, S., et al. “Three-Dimensional CMOS ICs Fabricated by Using Beam Recrystallization.” *IEEE Electron Device Letters*, vol. 4, no. 10, 1983, pp. 366-68, doi:10.1109/EDL.1983.25766.

3-Dimensional Gate Array with Vertically Stacked Dual SOI/CMOS Structure Fabricated by Beam Recrystallization. 1984, pp. 44-45.

Kleiner, M. B., et al. “Microelectronic Interconnects for 3-D Integration.” *IEEE Transactions on Components, Packaging, and Manufacturing Technology—Part B*, vol. 19, no. 4, 1996.

Knechtel, Johann, et al. “Large-Scale 3D Chips: Challenges and Solutions for Design Automation, Testing, and Trustworthy Integration.” *IPSJ Transactions on System LSI Design Methodology*, vol. 10, 2017, pp. 45-62.

Knechtel, Johann, Igor L. Markov, and Jens Lienig. “Assembling 2D Blocks into 3D Chips.” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 31, no. 2, 2012, pp. 228-41.

Koyanagi, Mitsuhiro, et al. “High-Density Through-Silicon Vias for 3-D LSIs.” *IEEE International Electron Devices Meeting (IEDM)*, 2008.

Larsen, Kenneth. “3D IC: Opportunities, Challenges, and Solutions.” *Semiconductor Engineering*, 28 Sept. 2021, semiengineering.com/3d-ic-opportunities-challenges-and-solutions/. Accessed 28 Jan. 2025.

Lau, John, et al. “Redistribution Layers (RDLs) for 2.5D/3D IC Integration.” *International Symposium on Microelectronics*, 2013, pp. 434-41, doi:10.4071/isom-2013-WA12.

Lau, John H. *Heterogeneous Integrations.* **Springer, 2019.

Lee, Hsien-Hsin S., and Krishnendu Chakrabarty. "Test Challenges for 3D Integrated Circuits." *IEEE Design & Test of Computers*, vol. 26, no. 5, 2009, pp. 26-35.

"Thermomechanical Stress-Aware Management for 3-D IC Designs." *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 25, no. 9, 2017, pp. 2678-82, doi:10.1109/TVLSI.2017.2707119.

Lobo, Savia. "Intel Unveils the First 3D Logic Chip Packaging Technology 'Foveros,' Powering Its New 10 nm Chips 'Sunny Cove'." *Packt Hub*, 13 Dec. 2018, hub.packtpub.com/intel-foveros-3d-logic-packaging/. Accessed 3 Mar. 2025.

Macchiolo, A., et al. "SLID-ICV Vertical Integration Technology for the ATLAS Pixel Upgrades." *Physics Procedia*, vol. 37, 2012, pp. 1009-15, doi:10.1016/j.phpro.2012.02.444.

Maestre Caro, A., et al. "Enabling Cu-Cu Connection in (Dual) Damascene Interconnects by Selective Deposition of Two Different SAM Molecules." *IEEE International Interconnect Technology Conference (IITC)*, 2011, pp. 1-3, doi:10.1109/IITC.2011.5940263.

Manners, David. "Samsung Makes 1 TB Flash eUFS Module." *Electronics Weekly*, 30 Jan. 2019, electronics-weekly.com. Accessed 3 Dec. 2024.

Mastroianni, Tony, and Todd Burkholder. "Taking 3D IC Heterogeneous Integration Mainstream." *3D InCites*, 26 Mar. 2025, 3dincites.com/2025/03/taking-3d-ic-heterogeneous-integration-mainstream/. Accessed 13 Nov. 2024.

Michallet, Jean-Eric. "CoolCube™: A True 3D-VLSI Alternative to Scaling." *3D InCites*, 24 Mar. 2014, 3dincites.com/2014/03/coolcube-3dvlsi-alternative/. Accessed 9 Jan. 2025.

Moskowitz, Sanford L. *Advanced Materials Innovation: Managing Global Technology in the 21st Century.* **Wiley, 2016.

National Institute of Standards and Technology (NIST). "CHIPS for America: Vision for a National Advanced Packaging Manufacturing Program." 20 Nov. 2023, nist.gov/news-events/news/2023/11/chips-america-releases-vision-approximately-3-billion-national-advanced. Accessed 26 Feb. 2025.

Palesko, Andrew. "The Cost of 3D ICs." *3D InCites Knowledge Portal*, 9 Jan. 2015, 3dincites.com/knowledge-portal/the-cost-of-3d-ics. Archived 9 Jan. 2015. Accessed 9 Mar. 2025.

Pandey, S., et al. "Assembly of a 3D Cellular Computer Using Folded E-Blocks." *Micromachines*, vol. 7, no. 5, 2016, article 78, doi:10.3390/mi7050078.

Patti, Robert S. "Three-Dimensional Integrated Circuits and the Future of System-on-Chip Designs." *Proceedings of the IEEE*, vol. 94, no. 6, 2006, pp. 1214-24.

Patterson, Alan. "$1.6 Billion CHIPS Subsidy Boosts Advanced Packaging in U.S." *EE Times*, 23 July 2024, eetimes.com/1-6-billion-chips-subsidy-boosts-advanced-packaging-in-u-s/. Accessed 11 Apr. 2025.

"Press Release: Ayar Labs Unveils World's First UCIe Optical Chiplet for AI Scale-Up Architectures." *Business Wire*, 31 Mar. 2025, businesswire.com/news/home/20250331044779/en/. Accessed 28 Oct. 2024.

Real World Technologies. "3D Integration: A Revolution in Design." 2 May 2007, realworldtech.com/3d-integration. Archived 22 Dec. 2010. Accessed 18 Dec. 2024.

Reif, Rafael, et al. "3-D Interconnects Using Cu Wafer Bonding: Technology and Applications." *Advanced Metallization Conference*, 2002, pp. 37-44.

Reiter, Herb. "Comparing Samsung's 3D NAND with Traditional 3D ICs." *3D InCites*, 16 Aug. 2013, 3dincites.com/2013/08/comparing-samsungs-3d-nand-with-traditional-3d-ics/. Accessed 13 Feb. 2025.

"Rochester Cube Points Way to More Powerful Chip Designs." *ScienceDaily*, 17 Sept. 2008, sciencedaily.com/releases/2008/09/080916151917.htm. Archived 17 May 2014. Accessed 13 Apr. 2025.

Savastiouk, Serge, Oleg Siniaguine, and Eric Korczynski. "Thru-Silicon Vias for 3D WLP." *Proceedings of the IEEE International Symposium on Advanced Packaging Materials*, 2000, pp. 206-07.

Sony Corporation. "Stacked CMOS Image Sensor Technology." *IEEE International Solid-State Circuits Conference (ISSCC) Technical Digest*, 2012.

Tanaka, Masayuki, et al. "3D Integration Technology and Heterogeneous Integration." *IEICE Transactions on Electronics*, vol. J94-C, no. 11, 2011, pp. 355-64.

Tezzaron Semiconductor. "3D-ICs and Integrated Circuit Security." 2008, tezzaron.com, PDF. Archived 7 Sept. 2008. Accessed 16 Jan. 2025.

"The Z Path Forward." Tezzaron Semiconductor, 2009, tezzaron.com/the-z-path-forward. Accessed 16 Jan. 2025.

Universal Chiplet Interconnect Express (UCIe) Consortium. "UCIe Consortium Announces Incorporation and New Board Members; Open for Membership." 2 Aug. 2022, nasdaq.com/press-release/ucie-universal-chiplet-interconnect-express-consortium-announces-incorporation. Accessed 13 Feb. 2025.

Virtex-7 2000T: 2D to 2.5D FPGA Innovation. Xilinx White Paper, 2011, xilinx.com/support/documentation/white_papers/wp380.pdf. Accessed 7 Feb. 2025.

Von Trapp, Françoise. "Monolithic 3D IC Heats Up at DATE 2015." *3D InCites*, 16 Mar. 2015, 3dincites.com/2015/03/monolithic-3d-ic-heats-up-at-date-2015/. Accessed 28 Nov. 2024.

Wang, Shuai, et al. "3D Integrated Circuit Cooling with Microfluidics." *Micromachines*, vol. 9, no. 6, 2018, article 287, doi:10.3390/mi9060287.

"What Is 3D Integration?" *3D InCites Knowledge Portal*, n.d., 3dincites.com/3d-incites-knowledge-portal/what-is-3d-integration/. Accessed 26 Apr. 2025.

Woo, Dong Hyuk, et al. "An Optimized 3D-Stacked Memory Architecture by Exploiting Excessive, High-Density TSV Bandwidth." *Proceedings of the 16th International Symposium on High-Performance Computer Architecture*, IEEE, Jan. 2010, pp. 429-40.

Xie, Jing, et al. "Design Space Exploration for 3D-Stacked Cache in Processor." *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 20, no. 3, 2012, pp. 527-40.

Yoon, Seung Wook, et al. "Fabrication and Packaging of Microbump Interconnections for 3D TSV." *IEEE International Conference on 3D System Integration*, 2009, pp. 1-5, doi:10.1109/3DIC.2009.5306554.

Zou, Qing, Emre Kursun, and Yuan Xie. "Thermomechanical Stress-Aware Management for 3D IC Designs." *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 25, no. 9, 2017, pp. 2678-82, doi:10.1109/TVLSI.2017.2707119.