INTERNATIONAL ROADMAP FOR DEVICES AND SYSTEMS™

2024 IRDS

EXECUTIVE PACKAGING TUTORIAL (EPT)

CONSIDERATIONS FOR HIGH PERFORMANCE COMPUTING AND ARTIFICIAL INTELLIGENCE

PART 1 – INTEGRATED CIRCUIT PACKAGING EVOLUTION

Table of Contents

List of Figures

List of Tables

DISCLAIMER

To provide realistic and tangible understanding of multi-chip module (MCM) technologies the Executive Packaging Tutorial (EPT) has collected information derived some of the most commonly available examples existing in the market to fulfill a realistic tutorial function. It is not the goal of the EPT to endorse or advertise any of these MCM approaches. Furthermore, since the assessment of these MCMs is done on the basis of publicly available information it is possible that the original suppliers may have additional knowledge and information that could change some of the assessments reported in the EPT. It is recommended that readers seeking additional insight on any of the reported MCM approaches and related subjects should be contacting the suppliers to obtain the most reliable information.

Finally, many of the MCM potential benefits and potential issues are reported on a case-by-case approach. These statements represent a technical opinion of the EPT and should not be constructed as benchmarking assessments since no experimental verification has been conducted in preparing this report. In addition, none of the assessments expressed for any specific MCM approach should be used as a benchmarking reference when comparing one MCM approach to another.

INTERNATIONAL ROADMAP FOR DEVICES AND SYSTEMS (IRDS)

EXECUTIVE PACKAGING TUTORIAL PART 1

Integrated Circuit Packaging Evolution

Overview

It is the purpose of this Executive Packaging Tutorial (EPT) to shed some light on the adoption, evolution and transformation of multi-chip modules (MCMs) in the era of high-performance computing (HPC) and artificial intelligence (AI) to help the reader to better comprehend the drivers and the roles of the fundamental building blocks of this technology revolution in a neutral and objective way.

This tutorial is providing a broad overview of the fast-evolving field of packaging. Until five years ago system on chip (SoC) products provided an adequate level of transistor integration necessary to realize a variety of multi-functional building blocks on a single chip meeting the requirements of the most advanced systems. The requirements of fast evolving HPC and AI systems have however outpaced the rate at which the required functionalities can be realized on a single die. This situation has re-energized packaging technologies developed as far back as 50 years ago for high performance but then prohibitively expensive systems. Presently, all the leading system integrators rely on packaging technologies to co-locate multiple dice in a single package (SiP) and multiple closely connected MCM in a single board. New packaging technologies have also been developed and deployed to accelerate an increased level of integration at the package and board levels. Many companies have developed proprietary nomenclatures to describe their own solutions. It is the purpose of this tutorial to provide a neutral overview of the field of advanced packaging.

Most of all the EPT consists in outlining step-by-step a major shift in packaging technology initially driven by HPC and presently accelerated by AI. Personal computers (PCs) and smart phones were the drivers of the electronics industry for the past 40 years. Their selling prices hovered in the range of thousands of dollars. Consequently, analysis of the bill of materials (BOM) documented that cost of the each of the integrated circuits spanned from a few dollars to tens of dollars. This limited allocation of dollars to each of the IC products consequently demanded to maintain the cost of packaging in the cents/pin range. To accomplish this goal packaging manufacturing was transferred oversees early on (i.e., mid-seventies) to mainly benefit from the low-cost of labor.

The cost of the packaging materials were also closely monitored and minimized. Packaging materials evolved from costly ceramic materials to plastic encapsulation to currently “organic” packaging substrates, just to mentioned a few. The downside of all these low-cost solutions resides in the implication that signal propagation delay resulted severely degraded right after the signal left the IC due to poor propagation properties of all the materials used for the in-package interconnections related to their intrinsic electrical properties. To partially circumvent this problem, multiple system in package technologies known at the time as multi-chip modules were developed about 30 years ago and used for very expensive HPC systems. New MCM applications were rediscovered in recent years driven by AI applications. SiP consists mainly in co-locating memory chips next to logic chips on the substate of a single package; by so doing overall space occupied by the components is minimized as compared to connecting several individually packaged chips co-located on the same board. This design results in signal propagation delays being reduced due to shorter lengths of interconnections (from cm to mm).

Discussions about further reducing the effect of the package organic substrate on interconnect delays are evolving towards the development of glass substates to provide better performance. This is perfectly reasonable based on a packaging driven solution, but this is where it is necessary to make a major departure from traditional approaches depending on the type of product.

The HPC and the AI worlds are presently undergoing a major transition from central processing unit (CPU)-driven HPC computational data centers to graphical processing unit (GPU)-driven AI data centers.

The basic building block of a datacenter is represented by the server “rack” whose price becomes the reference performance and cost benchmark to be compared to PCs and mobile phones prices. For example, a Blackwell-powered rack may cost up to several millions of dollars! Analyzing the price of a data center and going down the supply chain looking at the key components of a rack reveals that a Bella board may cost hundreds of thousands of dollars. A Blackwell chip may be priced up to $20, 000 to $70,000!

In this case also, any reduction in signal propagation delay remains essential and if a comparison is made to accomplish this goal between the ratio of traditional IC cost to packaging cost it becomes clear that tens and even hundreds of dollars can be allocated to the cost of an HPC or AI package.

With this realization it is necessary to abandon the world of low-cost driven packaging world and step into the HPC/AI-driven world of leading-edge expensive wafer manufacturing to get the right answer!

Recently, the concept of power delivery from the backside of an IC has been demonstrated (backside power distribution (BPD)) and several companies have announced plans to implement that technology. Additionally, the intention of distributing clock signals from the backside (backside clock (BC)) of a CPU or a GPU have been reported. This is realized by manufacturing these capabilities on a "partially sacrificial" wafer and then "thinning and pealing" the surface of this wafer where the innovative BPD and clock distribution circuitry is built. After these initial process steps, "bonding" the thinned down and peeled wafer surface to the backside of the microprocessor unit (MPU) or GPU wafer yields the desired structural configuration.

Furthermore, it appears that other active devices can be implemented in the sacrificial wafer since it is amenable to be produced with any leading-edge technology. With this realization it then appears clear that by optimizing the thickness of the thinned-down support wafer it is possible to extend its role to becoming the substrate of the whole SiP! This approach would also provide additional benefits. For instance, thermal compatibility between the upper chip and the chip bonded on the backside would be ensured since both are made of silicon. In addition, multiple levels of interconnections could be fabricated on this new silicon substrate to connect all the other chips located in the SiP with substantial reduction of signal propagation delays.

The cost of this complex substate produced by leading-edge silicon technology would likely escalate in the hundreds of dollars or more, but it should be remembered that the combined, integrated dice are selling for tens of thousands of dollars! This realization revisits the 60s to 90s era in the industry in which only system companies could afford the adoption of expensive flip chip packaging technology. This is because this cost disappeared in the overall cost of the system that was rewarded by very handsome selling prices in the ten thousands and hundred thousands of dollars.

Packaging History

Encapsulation of transistors was devised early in IC design and manufacturing to prevent any damage to the frail structure of transistors during handling. Connections from ICs to package were accomplished by bonding gold wires first to the aluminum bond pads of the IC by ultrasonic welding and then connecting these bond wires onto gold pads located inside the package case. Once the die was fully encapsulated within the package shell, elongated pins protruding from the bottom or side of the package and connected to the package bond pads were then inserted in a socket located on a printed circuit board (PCB). Electrical signals generated by individual transistors and later on by integrated circuit (IC) were connected via metallic traces located on the surface of the board to other components placed on the same PCB.

As time went by, the number of transistors and complexity of ICs kept on increasing and consequently die size increased. To accommodate for the increased number of signals exiting the ICs it became necessary to evolve the package to a dual in-line structure to accommodate for pins on two sides. Eventually the package evolved to a square format with pins on all four sides to maximize the number of available inputs/outputs for electrical and mechanical connections. During this evolution to more connections, the role of packaging remained limited to physical protection of ICs and as electrical conduits to the board and systems, thus mainly as a rather passive element. However, things drastically changed in the past 20 years as the role of packaging surged to the level of an active solution provider. The most dramatic example to illustrate this point occurred in 2011.

Qualcomm and Xilinx were both competing for market share in field-programmable gate array (FPGA) technology, these products were characterized by very large die sizes. Both companies relied on outside sources for the silicon wafers manufacturing. The die size of FPGA had kept on further increasing therefore progressively and drastically reducing the number of dice produced on each wafer due to both die size and die yields; conjunctively the 28 nm technology available on the market at the time was rather immature and it was not fully adequate to support both performance and yields required for FPGA. However, when technology by itself is not sufficient to solve a problem, it is time for creative engineering solutions. Xilinx made the insightful decision to split the large die into two dice. These were located side by side into the same package but electrically connected with each other so that operated as a single large die. It is well known that for a given number of defects the yields of smaller dice are higher than the yields of an equivalent die twice the area. Furthermore, the absolute amount of leakage and other deleterious effects were more manageable by partitioning the die into two dice. In the end this approach called 2.5D was successful and prevailed in the marketplace; for the first time the crucial role of advanced packaging stood up and forever changed the role of packaging from a passive participant to an active enabler. It was then said that packaging solutions could

provide contributions to the final product by providing introduction of a *"half technology node generation"* ahead of schedule.

It soon become clear that this first step in novel packaging technology was only the beginning of an intense integration between dice and packages in providing the most viable and timely system-oriented solutions.

In 2014 the various teams of the International Technology Roadmap for Semiconductors (ITRS) effort—realizing that many transformations in IC manufacturing were underway—warned that major changes were coming in the electronics industry, such as new IC/package roles. The ITRS identified that the relationship between systems, integrated circuits, and packages was becoming symbiotically intermingled going forward.

Due to this realization the ITRS reassembled the seventeen highly specialized technology working groups (TWGs) existing at the time into seven integrated focus teams (IFTs). The first reports of this new organization were published in the 2015 ITRS2.0. Among the newly created chapters a forward-looking chapter entitled "Heterogeneous Integration" was published and for the first time provided a novel view of the many integrated packaging and IC solutions coming in the next 15 years. It was expected that initially *homogeneous integration* of electronics components (i.e., chips based on similar silicon technologies) were going to populate the system in package world and then evolve into more complex ones in the subsequent 10-15 years. Within this projected timeline an era of truly heterogenous integration of ICs (created with fundamentally different technologies) was to occur starting with the year 2025; as such, this process is just beginning.

As predicted by the ITRS in 2015 and subsequently by the IRDS, the interaction and symbiosis between ICs and packages has continued to synchronically and homogenously increase and as a result multi-chip modules have now become fundamental building blocks in the manufacture of any high-performance electronics system. Many companies have created their own approach and *self-defined nomenclature*, and, even though some differences exist, it can be shown that fundamentally these are all variations on the same technology.

The Approach to the Executive Packaging Tutorial

This EPT has been assembled in three parts. Each one of them is essentially "self-contained." This approach allows the knowledgeable reader to advance to the section of their interest without necessarily reading several of the previous sections. This approach comes at a price since it creates some intrinsic redundancy but provides the benefit to the reader of "self-contained" sections. Therefore, the readers should be aware that this is not an editorial failure but it is motivated by a necessary level of redundancy.

This part of the EPT provides a brief overview of packaging technologies. However, it is not intended to represent an encyclopedia of packaging. The main purpose of Part 1 is to outline the features and evolution of packaging technologies leading to system in package for HPC and AI. However, this cannot be done ignoring the evolution of system on chip (SoC). A more extensive analysis of SoC is provided in the 2024 IRDS Executive Summary, and it is recommended that the two documents are evaluated side by side.

In order to make the content more easily readable many additional details are provided in the appended material for further insights. Details for process steps and examples of technologies are appended for elaboration purposes.

1. Relationship Between Transistor Switching Speed and Interconnect Delay

The relationship between transistor switching speed and interconnect delay is crucial in determining the overall performance of integrated circuits. As transistor switching speeds increase, the impact of interconnect delay becomes more significant, eventually becoming the limiting factor in circuit performance. Understanding this relationship and exploring solutions to minimize interconnect delay is essential for advancing performance of semiconductor technologies. Table 1 provides solutions to mitigate interconnect delay.

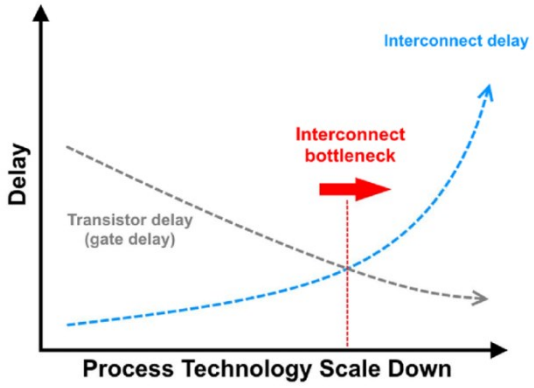


Figure 1. Qualitative comparison of transistor delay and interconnect delay

The Challenge of Signal Propagation Delay

1. **Transistor Switching Speed**
    - **Definition**—Transistor switching speed refers to the time it takes for a transistor to switch between its on (conducting) and off (non-conducting) states. This speed is often measured in terms of the gate delay or the rise and fall times of the transistor.
    - **Influence on Circuit Speed**—The faster a transistor can switch, the higher the potential clock frequency of a circuit, leading to faster data processing and overall performance.
2. **Interconnect Delay**
    - **Definition**— Interconnect delay is the time it takes for a signal to propagate through the metal wires (interconnects) that connect different components within an IC. This delay is influenced by the resistance (R) and capacitance (C) of the interconnect, often referred to as the RC delay.
    - **Significance**— As time went by transistors have become faster and therefore more numerous and overall longer interconnects have made interconnect delay the major bottleneck in achieving higher circuit speeds. In few words, even if transistors will continue to switch at increasingly higher speeds, the overall performance will be limited by the time it takes for signals to travel through the interconnects.
3. **Increasing Disparity**
    - **Scaling Challenges**—With transistor sizes shrinking (following Moore's Law), transistor switching speeds have increased. Concurrently, interconnects have scaled down also leading to both increased resistance (reduced cross section) and increased capacitance (closer line to line spacing) leading to higher RC values per unit length. This results in a growing disparity where interconnect delay becomes more pronounced compared to transistor switching speed.

Table 1. Solutions to Mitigate Interconnect Delay

Solution	Description
Material Innovations	• Low-k Dielectrics—Using low-k dielectric materials between interconnects reduces the capacitance, thereby reducing the RC delay. This has been a common approach in modern semiconductor manufacturing. • Copper Interconnects—Replacing aluminum with copper as the interconnect material reduces resistance, as copper has lower resistivity. This change helps to reduce the overall interconnect delay.
Advanced Interconnect Design	• 3D Integration and through-silicon vias—Three-dimensional (3D) integration using TSVs can significantly reduce the length of interconnects by allowing vertical stacking of circuits. This reduces the overall interconnect delay by shortening the distance signals need to travel. • Interconnect Spacing and Width—Optimizing the width and spacing of interconnects can help manage resistance and capacitance, reducing delay. Wider interconnects reduce resistance, while increased spacing reduces capacitance.
Signal Repeater Insertion	• Repeaters—Placing repeaters (buffers) at intervals along long interconnects can reduce delay by boosting the signal and mitigating the RC delay. This technique breaks a long interconnect into shorter segments, each with reduced delay. • Optimal Placement—The placement of repeaters must be optimized to balance the benefits of delay reduction with the power and area overhead they introduce.
Design Techniques	• Clock Tree Optimization—The clock distribution network, or clock tree, can be optimized to minimize skew and delay. Techniques like balanced clock trees or the use of low-skew clock buffers help ensure that all parts of the circuit receive the clock signal simultaneously. • Timing-Driven Layout—During the physical design phase, timing-driven layout tools can be used to optimize the placement and routing of components to minimize critical path delays, including those caused by interconnects.
Signal Integrity Management	• Shielding and Guard Rings—Implementing shielding (ground or power wires adjacent to signal lines) can reduce crosstalk and noise, improving signal integrity and reducing delay variability. • Differential Signaling—Using differential signaling, where two complementary signals are transmitted, can improve noise immunity and reduce the impact of delay variations.
On-Chip Networks and Architectures	• Network-on-Chip (NoC)—In complex multi-core processors, NoC architectures can be used to manage data flow between cores more efficiently, reducing the impact of interconnect delays. These networks are designed to handle the communication between cores with minimal latency. • Chiplet Architectures—Chiplet-based designs, where a system is divided into smaller, interconnected chiplets, can optimize interconnects within each chiplet and between them, reducing overall delay.
Emerging Technologies	• Optical Interconnects—Optical interconnects, which use light instead of electrical signals to transmit data, offer the potential for much higher speeds with lower delay over long distances. While still in development, they represent a promising solution for future high-performance ICs. • Carbon Nanotubes and Graphene—These materials offer lower resistance and higher electron mobility compared to traditional metals, potentially reducing interconnect delay if they can be reliably integrated into semiconductor manufacturing.

Take Away

The relationship between transistor switching speed and interconnect delay is a critical factor in modern IC design. While transistors have continued to get faster, interconnect delay has become the most significant challenge. Solutions

such as material innovations, advanced interconnect design, insertion of signal repeater, and the adoption of new technological and structural solutions are essential to mitigating interconnect delay and unlocking the full potential of fast-switching transistors. These approaches will help ensure that IC performance will keep pace with the increasing demands for better performance of modern HPC, AI and internet of things (IoT) communication systems.

The Challenge of Reducing Signal Propagation Delay

Signal delay analysis in ICs is a crucial aspect of circuit design, ensuring that signals propagate through the IC in a timely manner to meet performance and reliability requirements. An overview of key concepts and methods involved in signal delay analysis include types of delay, factors affecting signal delay, delay models, tools and techniques, and optimizing strategies. See Table 2.

Table 2 Considerations for Signal Delay Analysis

Key Concepts	Considerations
Types of Delay	• Propagation Delay—The time it takes for a signal to travel from the input to the output of a logic gate or circuit element. • Setup Time—The minimum time before the clock edge that the data input must be stable to ensure correct operation. • Hold Time—The minimum time after the clock edge that the data input must remain stable to ensure correct operation.
Factors Affecting Signal Delay	• Capacitance—The parasitic capacitance associated with the wiring and transistor gates affects delay. Increased capacitance slows down the signal. • Resistance—The resistance of the metal interconnects can also cause delays. Higher resistance leads to increased delay. • Transistor Characteristics—The size of transistors (width and length) and their threshold voltages impact how quickly they can switch, affecting delay. • Load—The load connected to the output of a gate or circuit affects the delay. Larger loads generally result in longer delays.
Delay Models	• RC Delay Model—This simple model approximates delay based on the resistance (R) and capacitance (C) of the interconnects. It's often used for quick estimations. • Logical Effort—This method involves analyzing how a gate's delay scales with the number of stages and the effort required to drive a load. It helps in optimizing circuit performance. • Elmore Delay Model—An extension of the RC model that provides a more accurate delay estimation by considering distributed capacitance.
Tools and Techniques	• SPICE Simulations—Used to simulate the detailed delay characteristics of ICs based on the actual circuit design and component models. • Static Timing Analysis (STA)—A method used to verify that the circuit meets timing requirements under all possible conditions. STA checks all possible paths in the circuit to ensure signals are properly synchronized. • Dynamic Timing Analysis—Examines timing under real operating conditions, including variability and environmental factors.
Optimization Strategies	• Sizing—Adjusting the width of transistors to balance delay and power consumption. • Pipelining—Breaking down long combinational paths into shorter stages separated by flip-flops to reduce delay. • Buffer Insertion—Adding buffers in long interconnect paths to reduce delay by driving the load more efficiently.
Critical Path Analysis	• Identifying the longest path through the circuit that determines the maximum clock frequency. Ensuring this path meets the timing requirements is crucial for reliable operation.
Impact of Technology Scaling	• As technology nodes shrink, delays in ICs can become more significant due to increased resistance, capacitance, and variations in manufacturing. Advanced analysis and optimization techniques are required to handle these challenges effectively.

Take Away Message

Signal delay analysis in IC design involves understanding and managing various factors that affect how quickly signals propagate through a circuit. By employing accurate models and sophisticated tools, designers can optimize performance and ensure that the IC meets all timing requirements.

Typical Values of Signal Propagation Delays

Signal propagation delays vary significantly depending on the material, distance, and specific technology used. Table 3 shows a general overview of typical signal propagation delays, as follows:

Table 3. Overview of typical signal propagation delays

Type	Propagation Delay	Details
Integrated Circuits	10 ps/mm to 100 ps/mm	In ICs, delays depend on the type of semiconductor technology (CMOS, BiCMOS, etc.), the complexity of the logic gates, and the interconnect length within the chip. Advanced nodes (e.g., 7nm, 5nm) tend to have lower delays due to shorter gate lengths and higher operating frequencies.
Packages	50 ps/mm to 150 ps/mm	The delay in a package is influenced by the type of package (e.g., wire-bonded, flip-chip, or ball grid array), the length of the interconnections, and the dielectric material used. For high-speed designs, package delay is a critical factor.
Printed Circuit Boards (PCBs)	140 ps/inch to 180 ps/inch (55 ps/cm to 70 ps/cm)	The delay on PCBs is influenced by the trace length, the dielectric material (FR-4, Rogers, etc.), and the trace width. High-speed PCBs are designed with controlled impedance to minimize delay and signal degradation.
Backplanes	140 ps/inch to 180 ps/inch (55 ps/cm to 70 ps/cm)	Backplanes are like PCBs but generally handle higher signal densities and longer distances. The propagation delay is usually like that of PCBs, but the length of traces in backplanes can lead to higher overall delay.
Trays and Racks (Cabling)	5 ns/m to 6 ns/m	In data center environments, signal propagation delays over cabling in trays and racks are influenced by the type of cable (e.g., coaxial, twisted pair, optical fiber). Optical cables have lower delays (~5 ns/m) compared to copper cables, and they also suffer less from signal attenuation over long distances.
Fiber Optics	~5 ns/m	Optical fibers are used in high-speed communication due to their low latency and high bandwidth. The delay is approximately 5 ns/m due to the speed of light in fiber being slightly slower than in a vacuum.

Why System on Chip?

Motivations and Drivers for System on Chip Development

The development and adoption of multiple and novel system on chip technologies have been historically driven by several motivations that focus on enhancing performance, reducing overall size of packaged die on PCBs, and improving efficiency. These motivations are key to meeting the demands of modern electronic devices, especially in areas like mobile computing, consumer electronics, and IoT where space is limited.

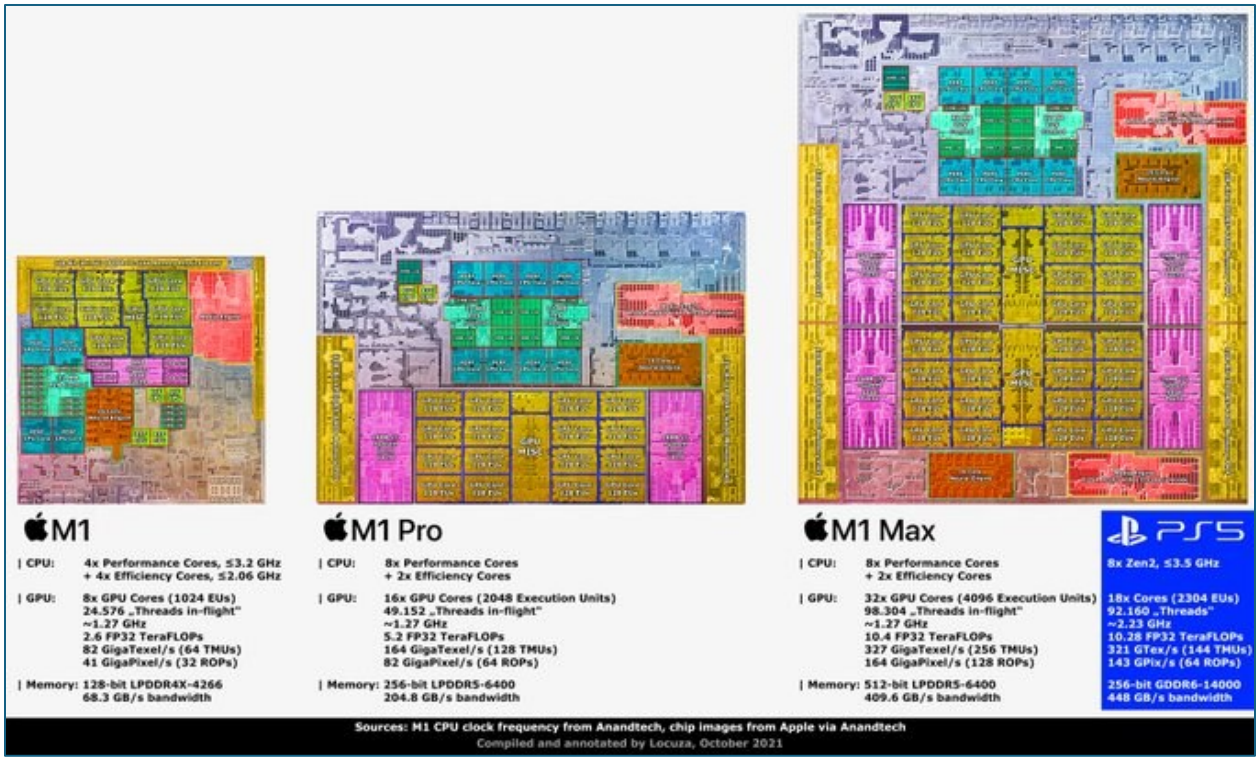


Figure 2. Examples of Apple System on Chip

Integration and Miniaturization

- **Increased Functionality**—SoCs integrate multiple components (such as CPUs, GPUs, memory, communication interfaces, and power management) onto a single chip.
    - This integration allows for greater functionality in a smaller form factor, which is essential for mobile devices, wearables, and other compact electronics.
- **Reduced Size and Weight**—By consolidating multiple discrete components into a single chip, SoCs significantly reduce the overall size and weight of electronic devices.
    - This is particularly important in industries where space is at a premium, such as aerospace, medical devices, and consumer electronics.

Power Efficiency

- **Lower Power Consumption**—SoCs are designed to optimize power usage by integrating power-efficient components and using advanced power management techniques.
    - This is critical for battery-powered devices, where energy efficiency directly impacts battery life.
- **Improved Thermal Management**—With fewer discrete components and shorter interconnections, SoCs generate less heat compared to multi-chip solutions.
    - This improves thermal performance, reducing the need for complex cooling solutions.

Performance Enhancements

- **Optimized Communication**—By integrating all components on a single chip, SoCs reduce the need for external interconnections, leading to faster data transfer rates and lower latency.
    - This improves overall system performance, especially in applications that require real-time processing.
- **Specialized Processing Units**—SoCs often include specialized processing units, such as AI accelerators, digital signal processors (DSPs), and graphics processing units.

- These units are optimized for specific tasks, enhancing the performance of applications like machine learning, image processing, and gaming.

Cost Reduction

- **Reduced Bill of Materials (BOM)**—By consolidating multiple functions into a single chip, SoCs reduce the number of components needed on the printed circuit board.
  - This leads to lower material costs and simplifies the supply chain.
- **Simplified Manufacturing**—SoCs reduce the complexity of the PCB design and assembly process. With fewer components to place and connect, the manufacturing process becomes more streamlined, leading to lower production costs and faster time-to-market.
- **Lower Packaging and Assembly Costs**—Integrating components into a single chip reduces the need for multiple packaging solutions, further cutting costs associated with assembly and testing.

Design Flexibility and Customization

- **Tailored Solutions**—SoCs can be customized to meet the specific needs of an application or market.
  - This flexibility allows companies to differentiate their products by integrating proprietary technologies or optimizing the chip for particular use cases.
- **Faster Innovation Cycles**—The ability to integrate new features and technologies into a single chip accelerates product development cycles.
  - This is crucial in fast-moving markets like consumer electronics, where staying ahead of competitors is vital.

Market Demands and Trends

- **Mobile and IoT Growth**—The explosion of mobile devices, IoT, and connected devices drives the demand for SoCs. These markets require compact, energy-efficient solutions that can deliver high performance in small, portable form factors.
- **AI and Edge Computing**—The increasing adoption of AI and edge computing applications has fueled the development of SoCs with integrated AI accelerators and other specialized processing units.
  - This enables devices to perform complex computations locally, reducing the need for cloud-based processing.

Cost Implications of SoC Development and Adoption

While SoCs offer significant advantages in terms of performance, power efficiency, and integration, their development and adoption come with certain cost implications that companies must consider. See Table 4.

Table 4. Implications of Development and Adoption Cost Type

Development and Adoption Cost Type	Implications
High Development Costs	• Complex Design Process—Designing an SoC is a highly complex and resource-intensive process. It involves significant investment in research and development (R&D), specialized design tools, and skilled engineers. The cost of developing an SoC can be substantial, particularly for cutting-edge technologies. • IP Licensing and Integration—SoCs often require the integration of third-party intellectual property (IP) blocks, such as cores for CPUs, GPUs, and communication interfaces. Licensing these IP blocks can add to the overall cost of SoC development.
Expensive Manufacturing	• Advanced Semiconductor Processes—SoCs are typically manufactured using advanced semiconductor processes, such as 7nm, 5nm, or even smaller nodes. These processes are expensive due to the complexity of

Development and Adoption Cost Type	Implications
	the technology, the need for specialized fabrication facilities, and lower yields associated with cutting-edge processes. • Testing and Validation—Comprehensive testing and validation are crucial for ensuring the reliability and performance of an SoC. This requires sophisticated testing equipment and methodologies, which can add to the manufacturing costs.
Economies of Scale	• High Initial Investment—The initial investment in SoC development and manufacturing is high. However, companies can achieve economies of scale by producing large volumes of SoCs, which helps spread the fixed costs over a larger number of units, reducing the per-unit cost. • Volume Production—To fully benefit from economies of scale, companies need to ensure high-volume production and strong demand for the SoC. Achieving this requires successful market penetration and widespread adoption of the devices that use the SoC.
Risk of Obsolescence	• Rapid Technological Advancements—The semiconductor industry evolves rapidly, with new process nodes, IP blocks, and design techniques continually emerging. SoCs developed today may become obsolete quickly, requiring continuous investment in new designs and technologies. • Market Shifts—Market demand for specific features or applications can change rapidly. An SoC tailored for a particular use case may face reduced demand if market trends shift, leading to potential losses if the SoC does not achieve expected sales volumes.
Long Development Cycles	• Time-to-Market Pressure—Developing an SoC takes time, and delays in the development process can lead to missed market opportunities. Companies need to carefully manage the development timeline to ensure that the SoC is ready for market introduction at the right time. • Design Iterations—The complexity of SoC design often requires multiple design iterations, each of which adds time and cost to the development process. Companies must balance the need for thorough testing and validation with the pressure to bring the product to market quickly.

Conclusion

The motivations and drivers for SoC development are rooted in the need for reducing the space occupied by the IC by means of higher integration while improving performance, providing better power efficiency, and enabling cost reduction in modern ICs.

While SoCs offer numerous advantages, their development and manufacturing come with significant costs. Companies must carefully consider these cost implications, including high development expenses, manufacturing costs, and the risks associated with rapid technological changes and market dynamics.

Despite these challenges, the benefits of SoCs often outweigh the costs, making them a crucial technology in the ongoing evolution of electronics.

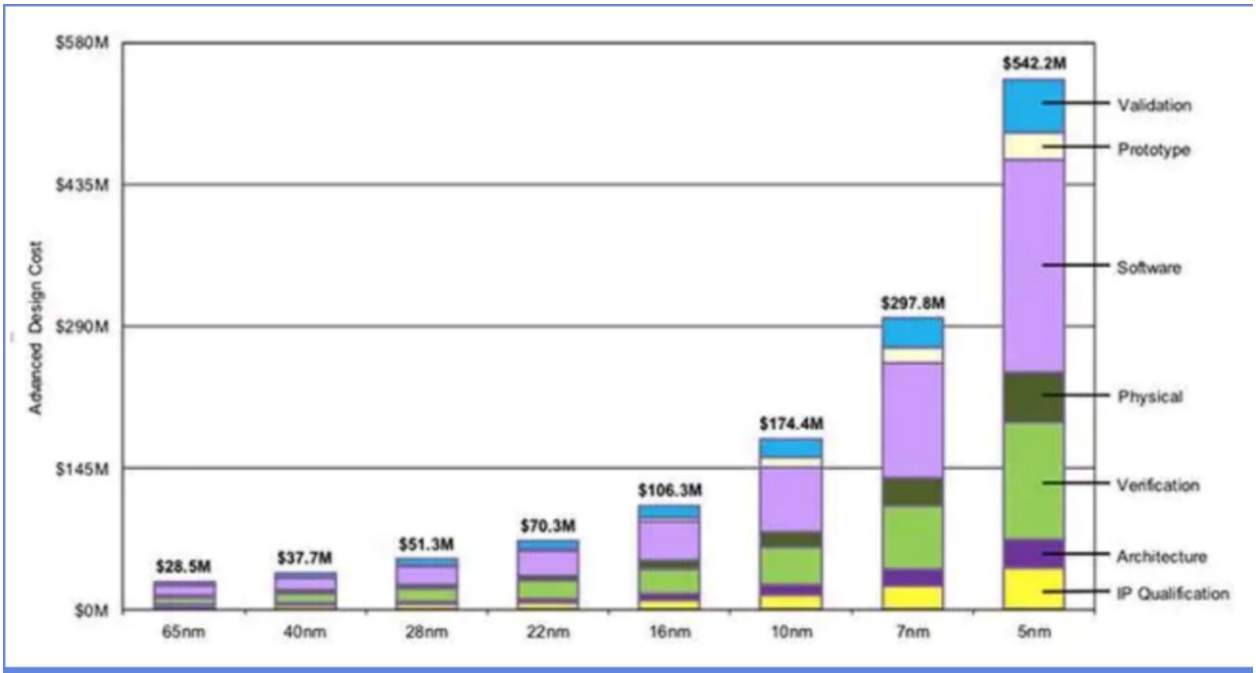


Figure 3. Total Cost Associated with SoC Product Design

Packaging Considerations of SoC and SiP

System-on-Chip packaging has evolved significantly since the concept of integrating multiple components onto a single chip first emerged. The following includes a brief overview of the history and development of SoC packaging:

Early Beginnings (1980s - 1990s)

- **Early Integration—**The concept of integrating multiple functions onto a single chip began in the 1980s with the rise of microcontrollers and early ASICs (Application-Specific Integrated Circuits). Initially, integration was limited to a few components, and packaging techniques were relatively simple, primarily using dual in-line packages (DIPs) and surface-mount packages.
- **Advancements in ASICs—**The 1990s saw the growth of more complex ASICs with higher levels of integration. These ASICs combined several functions onto a single chip and the associated packaging technology was progressively evolving, with packages like quad flat packages (QFPs) and ball grid arrays (BGAs) becoming more common.

Rise of SoCs (2000s)

Homogeneous Integration

- **Defining the SoC—**The term "System-on-Chip" became more prominent in the 2000s as semiconductor manufacturers began integrating entire systems, including processors, memory, I/O interfaces, and other components onto a single chip. This integration led to significant improvements in reduce PCB occupancy, improved performance and better power efficiency.
- **Advanced Packaging Techniques—**As SoCs became more complex, advanced packaging techniques were developed to manage the increased number of connections and reduce the overall footprint. Techniques in addition to wire bonding such as flip-chip bonding, and the use of advanced BGAs became common. Multi-chip modules were also used to combine multiple die in a single package.

Modern SoC Packaging in SiP (2010s - Present)

- **3D Packaging and Integration—**The 2010s saw the introduction of 3D packaging technologies, such as Through-Silicon Vias and stacked die configurations. These technologies allowed for greater integration density and improved performance by reducing signal latency and power consumption.
- **Advanced Substrates and Interposers—**The use of advanced substrates and interposers became more widespread. Technologies like Silicon Interposers and Embedded Multi-Die Interconnect Bridge (EMIB) provided high-bandwidth interconnects between different parts of the SoC or between the SoC and other components.
- **System-in-Package—**The concept of System-in-Package, which typically involves homogeneous integration of multiple logic and memory chips into a single package, became more common in the past 10 years. SiP allows also for the integration of various functionalities, such as RF components, sensors, and memory, alongside the main SoC.
- **Chiplet Architectures—**More recently, the industry has been moving towards chiplet-based architectures. This approach involves creating many small, modular chips (chiplets) that can be combined in a single package to create a versatile and scalable SoC. This method helps in managing design complexity and costs.

Future Trends

- **Heterogeneous Integration—**The future of SoC packaging is bound to eventually evolve from homogeneous integration and focus on heterogeneous integration, where different types of chips produced with multiple diverse technologies (e.g., III-V, GaN, Ge and other drastically different technologies) are integrated into a single package or module. This approach aims to further enhance performance, power efficiency, and functionality.
- **Advanced Materials and Techniques—**Ongoing research into new materials, such as advanced thermal interfaces and high-k dielectrics, as well as new packaging techniques, will continue to drive the evolution of SoC packaging, allowing for even higher integration densities and better performance.

Conclusion

Overall, the history of SoC packaging reflects the continual drive towards greater integration, higher performance, and improved efficiency, driven by the demands of modern electronic devices and applications. These requirements have prompted also the adoption of packaging technologies pioneered more than 30 years ago by system houses into consumers' applications.

First ETP Summary—System on Chip

SoCs presently integrate multiple homogeneous silicon-based components—such as CPUs, GPUs, memory, communication interfaces, and power management—onto a single chip. This integration allows for greater functionality in a smaller form factor. By consolidating multiple discrete components into a single chip, SoCs significantly reduce the overall size and weight of electronic devices. SoCs are designed to optimize power usage by integrating power-efficient components and using advanced power management techniques. This is critical for battery-powered devices, where energy efficiency directly impacts battery life. With fewer discrete components and shorter interconnections, SoCs generate less heat compared to multi-chip solutions. This improves thermal performance, reducing the need for complex cooling solutions.

But most of all by integrating multiple components on a single chip, SoCs reduce signal propagating delay and reduces the need for package-to-package interconnections, leading to faster data transfer rates and lower latency. This improves overall system performance, especially in applications that require real-time processing and minimal latency. These SoC accomplishments are nowadays spread across the electronics industry placing pressure on the packaging technology to concurrently evolve and contribute to accomplishing the same goals at the PCB level.

2. Electronic Packaging Overview

Electronics packaging is a critical aspect of designing and manufacturing electronic devices. It involves both the physical enclosure of the electronic components, and the methods used to protect and connect them. Table 5 presents the key areas in electronic packaging.

Table 5. Key areas involved in electronics packaging

Electronic Packaging Area	Description
Enclosures	These are the outer casings that protect electronic devices from physical damage, environmental factors, and electrical interference. Materials used can range from plastics and metals to specialized composites.
Thermal Management	Electronics often generate heat, so packaging needs to manage and dissipate this heat effectively. Heat sinks, thermal pads, and fans are common components used to ensure that electronic devices operate within safe temperature ranges.
Mechanical Protection	Packaging provides mechanical protection to delicate electronic components from shocks, vibrations, and other physical stresses during transportation and operation.
Electrical Connections	This includes connectors, soldering, and other methods used to establish electrical connections between different components of a device. Good design ensures reliable and efficient performance.
Signal Integrity	Proper packaging helps maintain the integrity of electrical signals by minimizing interference and cross-talk between components.
Environmental Protection	Packaging often includes measures to protect components from moisture, dust, and other environmental contaminants. This can involve sealing techniques, conformal coatings, and the use of protective enclosures.
Design and Aesthetics	The design of the packaging can also affect the user experience. It needs to be both functional and visually appealing, considering factors like ergonomics and branding.
Sustainability	There is increasing emphasis on using recyclable and environmentally friendly materials in electronics packaging to reduce waste and environmental impact.
Compliance and Safety	Packaging must adhere to various industry standards and regulations related to safety, electromagnetic compatibility (EMC), and other factors.

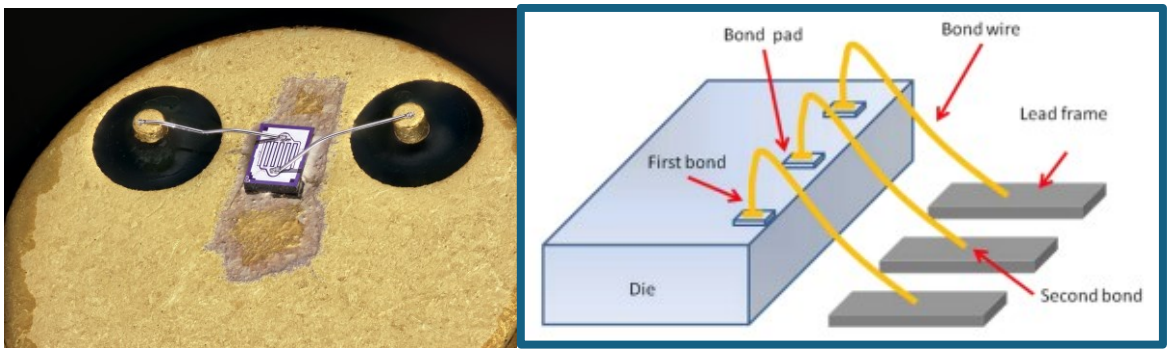
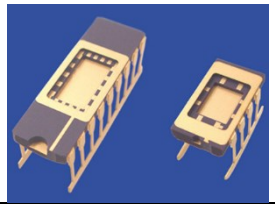
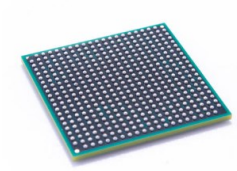
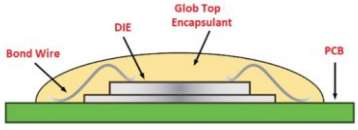


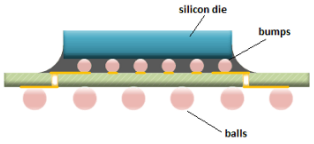
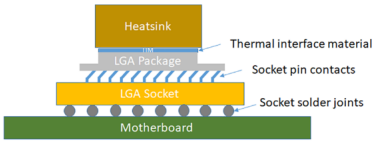
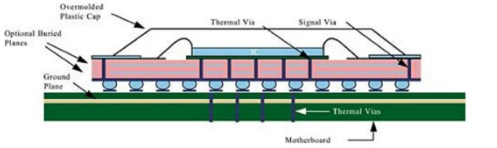
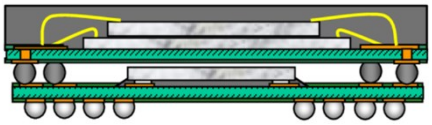
Figure 4. Early Wire Bond Packaging Solutions

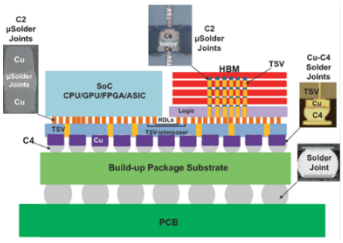
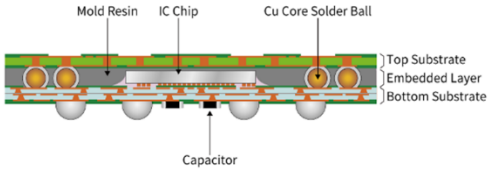
Summary of Packaging Technologies

Semiconductor packaging technologies have evolved significantly to meet the increasing demands for performance, miniaturization, and functionality in electronic devices. Table 6 presents a comprehensive list of major semiconductor packaging technologies and their main features.

Table 6. Major semiconductor packaging technologies and their main features

Packaging Technology	Main Features
Dual In-line Package Ceramic Packages A traditional package with two parallel rows of pins extending from the sides. Uses ceramic materials as the packaging substrate. 	• High Thermal Conductivity: Good for high-power and high-frequency applications. • High Reliability: Provides excellent protection against environmental factors. • Cost: Generally, more expensive than organic packages • Easy to Handle: Simple to insert and remove from sockets. • Through-Hole Mounting: Pins go through holes in the PCB. • Low Pin Count: Generally used for packages with a lower number of pins.
Surface-Mount Device (SMD) Components mounted directly onto the surface of the PCB without through-hole pins.  	• Compact Design: Reduces PCB space compared to through-hole designs. • Automated Assembly: Suitable for high-volume manufacturing. • Improved Performance: Shorter signal paths and better electrical performance.
Quad Flat Package (QFP) A package with leads extending from all four sides. 	• High Pin Count: Suitable for ICs with a high number of connections. • Surface-Mount Technology: Mounts directly onto the PCB. • Varied Pitch Sizes: Available in different lead pitches (distances between leads).
Ball Grid Array A package where the IC is connected to the PCB via an array of solder balls. (BGA) 	• High Pin Density: Accommodates a large number of connections in a compact area. • Improved Performance: Better electrical performance and thermal management. • Reflow Soldering: Solder balls are melted and solidified during reflow soldering.
Chip-on-Board (COB) Description: ICs are directly attached to the PCB, usually with wire bonding.  	• Cost-Effective: Reduces the need for a separate package. • Compact Design: Suitable for applications with space constraints. • Direct Electrical Connections: Wires connect the die directly to the PCB.
Flip-Chip (FC) The IC die is flipped over and mounted directly onto the substrate with solder bumps.	• High Performance: Shorter electrical paths improve performance.

Packaging Technology	Main Features
	• High Density: Allows for dense interconnects and higher pin counts. • Thermal Management: Direct heat dissipation through the substrate.
Land Grid Array (LGA) Description: Uses flat pads for connections. 	• Improved Reliability: Pads provide better alignment and less risk of solder joint defects. • Socket-Based: Often used in socketed applications for easy replacement. • High Pin Count: Suitable for high-density interconnections.
Ball Grid Array (BGA) Ball Grid Array (BGA) packaging is a type of surface-mount packaging used for integrated circuits. It's known for its ability to handle a high number of connections and its reliability in electronic devices. 	• Unlike traditional leaded packages, which have leads extending from the sides, BGAs have solder balls that are used for electrical connections and mechanical attachment to the printed circuit board.
Package-on-Package (PoP) Stacking of multiple packages on top of each other. 	• Space Savings: Allows for vertical integration of components. • High Integration: Combines memory and logic devices in a single package. • Thermal Management: May require additional heat dissipation solutions.
System-in-Package Integration of multiple ICs and passive components into a single package. 	• Compact Design: Integrates various functionalities into a single package. • Reduced Interconnect Lengths: Improves performance and reliability. • Flexibility: Allows for integration of heterogeneous components.

Packaging Technology	Main Features
3D Integrated Circuits (3D IC) Stacking multiple ICs vertically and connecting them with through-silicon vias. 	• High Bandwidth: Shorter interconnects enhance data transfer speeds. • Space Efficiency: Reduces footprint compared to traditional 2D layouts. • Improved Thermal Management: Requires advanced thermal solutions due to heat accumulation.
Package Substrate with Embedded Components Integrates components such as passive devices within the package substrate itself. 	• Reduced Size: Embeds components to save space. • Enhanced Performance: Shorter signal paths and improved signal integrity. • Complex Fabrication: Requires sophisticated manufacturing processes.
Flex Circuits Uses flexible materials like polyimide to create circuits that can bend and conform to various shapes. 	• Flexibility: Allows for designs in constrained spaces and dynamic applications. • Lightweight: Reduces overall weight compared to rigid PCBs. • Complex Designs: Suitable for applications requiring flexible or wearable electronics.

Conclusion

Each packaging technology offers unique benefits and is suited to different applications based on factors such as performance requirements, space constraints, and cost considerations. From traditional packages like DIP and QFP to advanced technologies like 3D IC and SiP, the choice of packaging technology plays a crucial role in determining the overall functionality, reliability, and efficiency of semiconductor devices.

Second ETP Summary—Packaging Evolution

In the past 50 years a multitude of packages were developed for different applications. As the number of transistors in a single die kept on increasing at Moore's Law pace it became necessary to increase the number of connections from the package to the outside world. This led to the evolution from the initial dual in line packages to multiple families of squarely shaped packages. Furthermore, packaging pins evolved and gave way to several forms of sturdily ball shaped connectors improving integrity and evolving also into a grid of connectors arranged across the whole area of the package to further increase the total number of possible connections to the outside. Packaging technology had also to evolve from single die to multiple dice packaged onto a single substate (SiP). This trend is continuing going

forward and moving from 2D arrangements to 3D arrangements. These subjects will be discussed in more depth in the subsequent chapters.

3. Packaging Substrates Overview

Substrate Materials

Substrate materials for semiconductor packaging play a crucial role in the overall performance, reliability, and cost of the final package. These materials provide the physical and electrical interface between the semiconductor die, the package and the printed circuit board. Below is an overview of the commonly used substrate materials and Table 7 presents substrate types and applications. Table 8 shows main building blocks of substrate manufacturing process.

Types of Substrates

Organic—Typically made from materials like fiberglass-reinforced epoxy laminates (e.g., FR-4) or high-performance polymers like polyimide. These are common in various packaging types like BGA (Ball Grid Array) and QFN (Quad Flat No-lead).

Ceramic—Made from materials like alumina (Al_2O_3) or aluminum nitride (AlN), offering excellent thermal and electrical properties, ideal for high-reliability applications like aerospace or automotive.

Glass, Metal and Silicon—Used in advanced packaging technologies like fan-out wafer-level packaging (FOWLP) and 2.5D/3D ICs.

Table 7. Substrate Types and Applications

Type of Organic Substrate	Description	Application
BT (Bismaleimide Triazine) Resin	BT resin is widely used in ball grid array (BGA) packages due to its excellent mechanical and thermal properties. It has good electrical insulation and is suitable for high-density interconnects.	BGAs, chip-scale packages (CSPs).
FR-4 (Flame Retardant 4)	FR-4 is a common, cost-effective material used in PCBs and some package substrates. It's an epoxy resin reinforced with woven glass fabric. While not the best for high-frequency applications, it's widely used due to its low cost and adequate performance.	Low to medium-performance packages.
Ajinomoto Build-up Film (ABF)	ABF is a type of resin-coated copper substrate used for high-density packaging in advanced ICs. It allows for finer line width and spacing, which is critical for advanced packaging technologies like flip-chip.	High-performance flip-chip packages, CPUs, GPUs.
Type of Ceramic Substrate	**Description**	**Application**
Alumina (Al_2O_3)	Alumina is a widely used ceramic substrate material known for its excellent electrical insulation, thermal conductivity, and mechanical strength. It's suitable for high-reliability applications.	High-power devices, military and aerospace applications.
Aluminum Nitride (AlN)	AlN offers superior thermal conductivity compared to alumina, making it ideal for applications requiring efficient heat dissipation. It also has good electrical insulation properties.	High-power and high-frequency devices, light-emitting diodes (LEDs).

Beryllium Oxide (BeO)	BeO has even higher thermal conductivity than AlN but is less commonly used due to its toxicity and handling difficulties. It's used in very specialized applications.	Extreme high-power applications.
Silicon Carbide (SiC)	SiC substrates are used in high-temperature and high-power applications due to their excellent thermal and electrical properties. They are often used in conjunction with SiC or GaN devices.	Power electronics, RF devices.
Type of Glass, Metal or Silicon Substrate	**Description**	**Application**
Glass	Glass substrates are emerging as a promising material for advanced packaging due to their excellent electrical insulation, low loss at high frequencies, and potential for fine-pitch interconnects. They offer good dimensional stability and are compatible with through-glass vias (TGVs).	Advanced packaging, RF, and photonic devices.
Metal Core--Copper Core	Copper core substrates are used in packages where high thermal dissipation is required. The copper core provides a path for heat to be efficiently spread and dissipated.	High-power devices, LEDs.
Metal Core--Aluminum Core	Like copper core, aluminum core substrates are used for thermal management but are lighter and less expensive. They offer good thermal conductivity and are commonly used in LED packaging.	LED packages, power electronics.
Silicon	Silicon substrates are used in advanced packaging techniques like silicon interposers in 2.5D and 3D ICs. They allow for high-density interconnects and integration of multiple chips in a single package.	2.5D/3D ICs, high-performance computing, and AI chips.

Table 8. Main Building Blocks of Substrate Manufacturing Process

Process Step	Description
Material Preparation	• Organic Substrates: Involve layering sheets of fiberglass with epoxy resin, followed by lamination under heat and pressure. • Ceramic Substrates: Start with ceramic powders that are mixed, cast, and sintered to form dense, durable layers. • Silicon/Glass Substrates: Sliced from silicon or glass wafers and prepared for further processing.
Circuit Patterning	• Photoresist Coating: A layer of photoresist is applied to the substrate surface. • Photolithography: UV light is used to pattern the photoresist based on the desired circuit design. • Etching: Unprotected areas of the substrate are etched away to form the circuit pathways.
Metallization	• Electroplating or Sputtering: To deposit metal layers, usually copper, which form the conductive traces.

Process Step	Description
	• Surface Finishing: May include processes like nickel-gold plating to improve solderability and corrosion resistance.
Layer Stacking (for Multi-layer Substrates):	• Via Formation: Through-holes (vias) are drilled and plated to connect different layers. • Lamination: Multiple layers are aligned and laminated together under heat and pressure.
Inspection and Testing:	• Includes electrical testing, optical inspection, and possibly X-ray inspection to ensure quality and reliability.
Field-Related Aspects:	• Field-Effect Transistor (FET) Impact: In semiconductor packaging, managing electric fields is crucial, especially in high-frequency applications. Substrate materials and design must minimize parasitic capacitance and inductance. • Electromagnetic Interference (EMI): Substrate design and material selection must account for shielding and minimizing EMI, especially in densely packed environments.

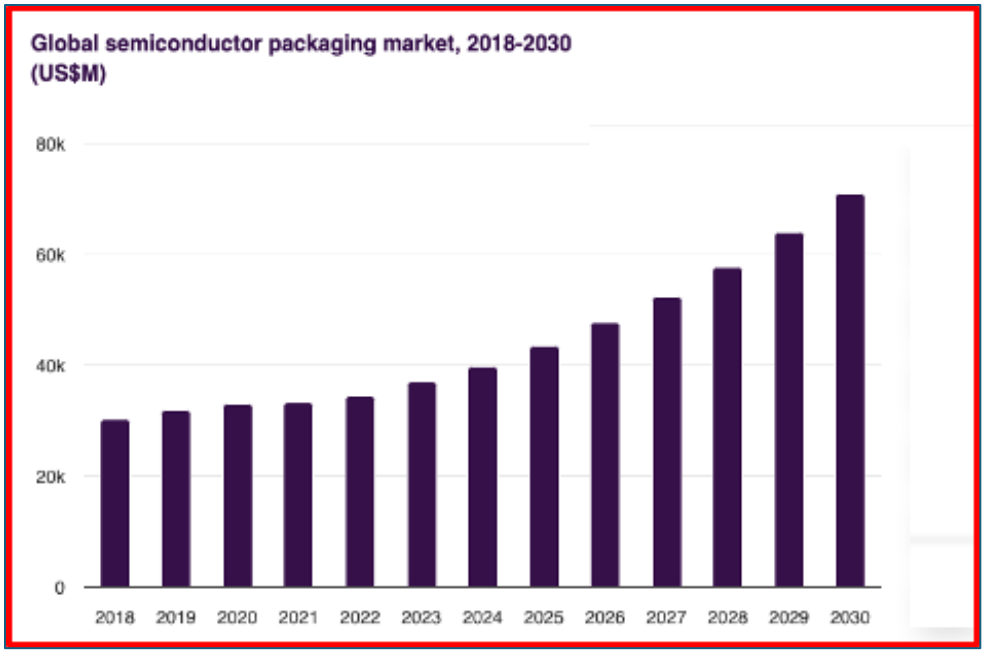
Take-away

Each substrate material has its strengths and weaknesses, making them suitable for different applications based on performance requirements, cost, thermal management needs, and electrical characteristics. What used to be a well-established packaging process with well-defined materials is now evolving towards more sophisticated structures and materials.

Factors Influencing Substrate Size

The size of a packaging substrate is a balance between the needs of the application, the capabilities of the manufacturing process, and the performance requirements of the final device. Appendix A details substrate sizes for different types of packaging substrates.

- Application Requirements—Higher power or more complex systems typically require larger substrates.
- Manufacturing Capabilities—The size is limited by the precision of the manufacturing processes and the equipment used.
- Thermal and Mechanical Considerations—Larger substrates may require additional thermal management and mechanical support.



Source: horizon-query@grandviewresearch.com

Figure 2. Overview of packaging market

Manufacturing Process of Substrate Packaging—Detailed Steps

The manufacturing process flow for substrate packaging, which is essential in semiconductor packaging, involves multiple steps that ensure the reliable integration of ICs with the substrate. The substrate serves as the foundation that provides mechanical support, electrical connections, and thermal management for the IC. Appendix A details a typical manufacturing process flow for substrate packaging.

Evaluation of Electronic Packaging Substrate Alternatives

When it comes to IC packaging, the choice between organic and glass substrates plays a critical role in determining the performance, reliability, and cost of the final product. Each type of substrate offers distinct advantages and disadvantages, making them suitable for different applications.

Applications

- **Organic Substrates**—Commonly used in consumer electronics, automotive electronics, and general-purpose ICs. They are suitable for most applications where cost and flexibility are more important than the absolute highest performance.
- **Glass Substrates**—Often used in high-performance applications such as RF and microwave ICs, high-speed digital circuits, and specialized high-reliability applications. They are chosen for applications requiring superior thermal and electrical performance.

Advantages and disadvantages for organic and glass substrates are compared in the following Table 9.

Table 9. Advantages and Disadvantages of Substrate Types

Substrate Type	Advantages	Disadvantages
Organic substrates are widely used in IC packaging due to their flexibility, cost-effectiveness, and ease of processing. They are typically made from materials such as epoxy resins or polyimide films and often include fillers to improve performance.	• Cost: Organic substrates are generally less expensive than glass substrates. They are also easier and cheaper to process, making them cost-effective for high-volume production. • Flexibility: These substrates can be easily fabricated into various shapes and sizes, making them versatile for different packaging needs. • Thermal Expansion Matching: Organic substrates have thermal expansion coefficients that can be matched with the silicon chip, reducing thermal stress during operation. • Electrical Performance: For many applications, organic substrates provide adequate electrical performance, including reasonable signal integrity and impedance control.	• Thermal Performance: Organic substrates typically have lower thermal conductivity compared to glass, which can be a disadvantage for high-power or high-frequency applications. • Moisture Sensitivity: Organic materials are more susceptible to moisture absorption, which can affect reliability and performance over time. • Signal Integrity: At very high frequencies, organic substrates may not provide as good signal integrity or impedance control as glass substrates.
Glass substrates are used in some advanced IC packaging applications, particularly where high performance and reliability are crucial. Glass is a rigid, thermally stable material that can provide excellent performance in specific scenarios.	• Thermal Performance: Glass has excellent thermal conductivity and thermal stability, which helps in dissipating heat more effectively compared to organic substrates. • Electrical Performance: Glass substrates offer superior signal integrity and impedance control, making them suitable for high-frequency and high-speed applications. • Moisture Resistance: Glass substrates are impervious to moisture, which improves long-term reliability and reduces the risk of degradation over time. • Flatness and Precision: Glass provides a very flat and precise surface, which is beneficial for fine-pitch interconnects and high-density packaging.	• Cost: Glass substrates are generally more expensive than organic substrates due to the higher cost of materials and more complex processing. • Processing Complexity: The fabrication and handling of glass substrates can be more complex and less flexible compared to organic substrates. • Thermal Expansion Mismatch: Glass has a different thermal expansion coefficient compared to silicon, which can lead to thermal stress issues if not properly managed.

Conclusion

The choice between organic and glass substrates depends on the specific requirements of the application:

- **For cost-sensitive, high-volume applications with moderate performance needs**—Organic substrates are generally preferred due to their lower cost and flexibility.
- **For high-performance, high-reliability applications where thermal and electrical properties are critical**—Glass substrates offer superior performance, albeit at a higher cost and with increased processing complexity.

Each substrate type has its place in the IC packaging landscape, and the selection decision should align with the cost and performance.

Third EPT Summary—Substrate Materials and Manufacturing Processes

Several materials are used as substrates in packaging. Each substrate technology is developed for specific market applications. The packaging driver remains to minimize cost while satisfying the basic product requirements and this explains while a single unified substrate solution has not emerged. The typical manufacturing process flow has been reviewed. Substrate materials and sizes were also summarized. Cost remains a key requirement. The market outlook remains positive.

Examples of some packaging technologies aimed at applications where many signals needed to be transferred from the IC to the outside world via packaging solutions.

4. Grid Array Technologies

Design of ICs has increasingly provided more functionalities, and it has become progressively more complex. As a result the number of pins required in a package to accommodate incoming and extracting electrical signals has exponentially increased causing a dramatic increase in package size and a drastic structural transformation in packages. This chapter provides an overview of the different packaging solutions developed to satisfy these needs.

Land Grid Array

Land Grid Array (LGA) is a type of surface-mount package where the IC is connected to the PCB via flat pads on the underside of the package, rather than solder balls like in Ball Grid Arrays (BGAs). Figure 4 is an illustration of an LGA. Refer to Appendix B for details of the LGA fabrication process.

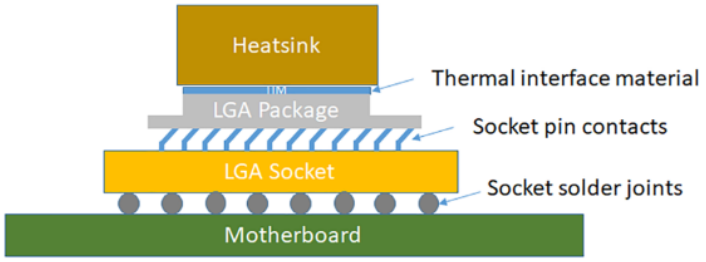


Figure 5. Example of LGA package mounted on motherboard and including heatsink

Land Grid Array Applications

LGA packaging is used in various applications where high pin density, robust mechanical connection, and precise alignment are required. Here are some key applications below. Table 10 lists advantages and disadvantages of LGA packaging.

- **High-Performance Processors**:
  - **Example**: Intel's high-performance CPUs.
  - **Reason**: LGA packages are used for processors where high-density connections and thermal performance are critical. The flat pads of LGA allow for good thermal dissipation and high pin counts.

- **Networking and Communication Equipment**:
  - **Example**: Routers, switches, and network interface cards (NICs).
  - **Reason**: These devices require reliable and high-speed connections, which LGA packaging provides. The LGA package's robust connection helps ensure signal integrity and performance.
- **Server and Data Center Components**:
  - **Example**: Server motherboards and data center processors.
  - **Reason**: Servers often use LGA packages due to their ability to handle high pin counts and the need for reliable, high-performance components.
- **Consumer Electronics**:
  - **Example**: High-end consumer electronics that require compact and reliable components.
  - **Reason**: LGA packages offer a compact design and reliable electrical and mechanical connections, making them suitable for advanced consumer electronics.
- **Automotive Electronics**:
  - **Example**: Advanced driver-assistance systems (ADAS) and infotainment systems.
  - **Reason**: Automotive applications benefit from the mechanical reliability and high pin density of LGA packages, which are important for the demanding conditions in automotive environments.
- **Test and Measurement Equipment**:
  - **Example**: Oscilloscopes, analyzers, and other precision measurement tools.
  - **Reason**: These devices require high-performance ICs that often use LGA packaging to ensure precise measurements and reliable operation.

Table 10. Advantages and Disadvantages of LGA Packaging

Advantage	Disadvantage
• High Pin Density: Allows for many connections in a compact footprint. • Improved Mechanical Stability: Flat pads provide a strong and reliable mechanical connection to the PCB. • Reduced Risk of Solder Joint Failure: The absence of solder balls reduces the risk of solder joint defects. • Enhanced Thermal Performance: Good thermal management due to the large contact area.	• Inspection Challenges: Solder joints are not visible after soldering, requiring advanced inspection methods like X-ray inspection. • Rework Difficulty: Reworking or replacing LGA packages can be more challenging compared to other package types due to the flat pad design.

Overall, LGA packaging is a versatile and reliable option for high-performance and high-density electronic applications, offering benefits in both electrical and mechanical performance.

Ball Grid Array

Ball Grid Array (BGA) packaging is a type of surface-mount packaging known for its ability to handle a high number of connections and its reliability in electronic devices. BGAs make electrical connections with ICs using an array of solder balls on the underside of the package. Unlike traditional leaded packages, which have leads extending from the sides, BGAs have solder balls that are used for electrical connections and mechanical attachment to a PCB. Figure 5 is an illustration of a BGA. Appendix C details the BGA manufacturing process. Table 11 lists advantages and disadvantages of BGA packaging.

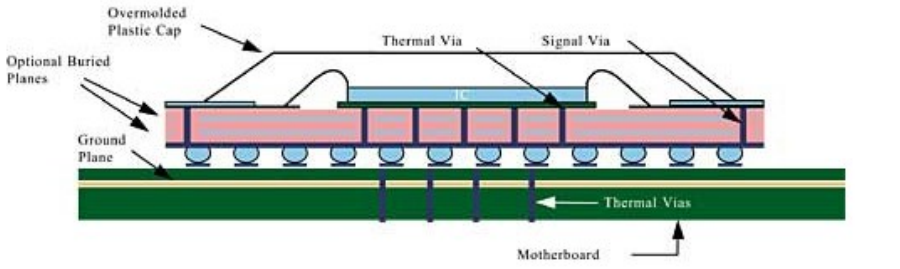


Figure 3. Example of BGA package

Key Features

- **Solder Balls—**BGA packages use an array of small solder balls arranged in a grid pattern. These balls are used to connect the package to the PCB. The size and spacing of the balls depend on the pitch of the BGA.
- **High Pin Count—**BGAs are designed to accommodate a large number of pins (or balls) within a relatively small area. This makes them suitable for high-density applications where space is limited.
- **Enhanced Performance—**The shorter electrical paths between the IC and the PCB reduce signal transmission delays and power loss, improving overall performance and signal integrity.
- **Thermal and Mechanical Reliability—**BGA packages typically offer better thermal performance and mechanical strength compared to traditional packages. The large area of the solder balls helps dissipate heat more effectively, and the package's solid connection to the PCB provides better mechanical stability.

Types of BGA Packages

- **Standard BGA (SBGA)—**The basic form of BGA packaging with a regular grid of solder balls. It is commonly used in various applications.
- **Fine-Pitch BGA (FBGA)—**A type of BGA with a smaller ball pitch (spacing between balls), allowing for higher pin density. This is used in applications where space is at a premium.
- **Chip-Scale BGA (CSBGA)—**This type of BGA has a package size that is nearly the same as the die size. It is used for very compact designs.
- **Ball Grid Array with Heat Spreader (BGA-HS) —**Includes a heat spreader or heat sink to help manage thermal dissipation in high-power applications.
- **Dual-Row BGA (DR-BGA)—**Features two rows of solder balls, allowing for a higher pin count in a smaller footprint.

Table 11. Ball Grid Array Packaging Advantages and Disadvantages

Advantage	Disadvantage
• Improved Electrical Performance: Shorter signal paths reduce signal delay and power loss, improving overall electrical performance. • Higher Pin Density: Allows for a greater number of connections within a smaller package footprint, making it suitable for high-density applications. • Better Thermal Management: Larger surface area for heat dissipation compared to traditional packages. • Mechanical Reliability: Stronger mechanical connection to the PCB, reducing the risk of solder joint failures.	• Inspection and Repair: Visual inspection of solder joints can be difficult due to the package's design. Automated optical inspection (AOI) and X-ray inspection are often required. • Cost: BGA packages can be more expensive to manufacture and assemble compared to simpler packages due to the complexity of the process and materials used. • Rework Difficulty: Reworking or replacing a BGA package is more challenging compared to other types of packages, which can impact repair and refurbishment processes.

BGA Applications

BGAs are widely used in various applications, including:

- **Consumer Electronics**—Smartphones, tablets, and laptops.
- **Networking Equipment**—Routers, switches, and other networking devices.
- **Automotive Electronics**—Engine control units, infotainment systems, and sensors.
- **Medical Devices**—Diagnostic equipment and medical imaging systems.

Overall, BGA packaging provides a high-performance, reliable solution for modern electronic devices, accommodating complex and high-density circuits with improved thermal and mechanical characteristics.

Ball Grid Array Connectivity

In a Ball Grid Array (BGA) package, the electrical connection between the die and the PCB (Printed Circuit Board) is achieved through a combination of bonding methods and solder balls. Appendix C details the BGA Electrical Connection Process.

Electrical Connection Details

- **Signal Pathways—**Electrical signals travel from the die through the internal routing of the package substrate, reaching the solder balls. From the solder balls, signals are transmitted to the PCB and connected to other components or circuitry.
- **Thermal and Electrical Performance—**The short distance between the die and the PCB, facilitated by the BGA's solder balls, contributes to better electrical performance by reducing signal path lengths and improving thermal dissipation.
- **Ball Grid Array Design—**The arrangement of solder balls in a grid pattern allows for a high density of connections in a compact area. The ball pitch (the distance between the centers of adjacent solder balls) determines the number of connections that can be accommodated within the package footprint.

Conclusion

The electrical connection between the die and the PCB in a BGA package is established through a combination of die bonding (either with wire bonding or flip-chip bonding), a package substrate that routes signals to the solder balls, and the solder balls themselves, which create the final connection to the PCB. This method provides high-density, reliable, and efficient electrical and mechanical connections suitable for a wide range of electronic applications.

Fourth EPT Summary—Grid Arrays

Two of the most popular packaging families were reviewed. These represent the early attempts to distributing the electrical output signals generated by the die into a grid array to maximize the numbers of outputs for a given package size but this approach still presented many drawbacks. In the next chapter the most successful approach to maximize outside distribution of multiple signals is reviewed.

5. Flip Chip Technology

Flip Chip Technology is a fundamental packaging technology enabling nowadays the most advanced HPC and AI systems. Many companies have created different nomenclatures to make this technology “their own”. The following paragraphs will extensively cover all the aspects of this technology.

Flip Chip History

Controlled Collapse Chip Connection (C4), commonly known as "C4 bump" or "flip-chip" technology, is a method of semiconductor device packaging that allows for the direct electrical connection of a semiconductor die to a substrate or circuit board. This technology represented a significant advancement over traditional wire bonding methods, enabling more compact, reliable, and high-performance connections in electronic devices. Here's an overview of the history and early applications of C4 technology:

History of Controlled Collapse Chip Connection (C4)

- **Development at IBM (1960s)**:
    1. **Inventor—**C4 technology was developed at IBM in the early 1960s by Dr. Leo Bakker, who worked on improving the packaging and interconnect methods for integrated circuits.
    2. **Need for Innovation—**As transistor counts increased and circuits became more complex, the limitations of traditional wire bonding became apparent. Wire bonds were prone to damage, could only connect to the edges of the die, and had significant limitations in terms of density and performance.
- **Introduction of Flip-Chip**:
    1. **Concept—**The core idea behind C4 technology was to "flip" the semiconductor die upside down and directly bond it to the substrate using small solder bumps made of Pb and Sn. This allowed for connections to be made across the entire surface of the die rather than just the edges.
    2. **Controlled Collapse—**The term "controlled collapse" refers to the way the solder bumps would collapse in a controlled manner when heated, ensuring reliable electrical connections and proper alignment between the die and the substrate.
- **Key Innovations**:
    1. **Solder Bumps—**Small, precisely placed solder bumps were used to create the electrical connections between the die and the substrate. These bumps were deposited on the die and then reflowed (melted) to create a bond.
    2. **Underfill Material—**To improve the mechanical strength and reliability of the connection, an underfill material was applied between the die and the substrate. This material helped distribute stress and protect the solder joints from thermal and mechanical stresses.

Early Applications of C4 Technology:

Mainframe Computers (IBM System/360 and System/370)

First Commercial Use—C4 technology was first used in IBM's mainframe computers, particularly in the System/360, introduced in 1964, and later in the System/370.

Performance and Reliability—The adoption of C4 in these systems allowed IBM to create more powerful and reliable computers with higher transistor densities. The technology also contributed to improved signal integrity and reduced parasitic inductance compared to wire bonding.

High-Performance Processors

IBM 3081 Processor—Another early application of C4 was in the IBM 3081 processor, introduced in the early 1980s. The use of C4 technology enabled higher clock speeds and better thermal management, which were crucial for the performance demands of this high-end processor.

Impact on Processor Design—C4 technology allowed for more complex and powerful processor designs, laying the groundwork for future advancements in microprocessor technology.

Military and Aerospace Applications

Rugged Environments—Due to its robustness and reliability, C4 technology was quickly adopted in military and aerospace applications. The ability to create dense, reliable connections in a compact form factor was highly valued in environments where space, weight, and durability were critical.

Radar and Communication Systems—C4 was used in radar systems, communication devices, and other critical electronic systems that required high performance and reliability under extreme conditions.

Advanced Packaging in the Semiconductor Industry

Transition to Consumer Electronics—Over time, as the semiconductor industry evolved, C4 technology became more widespread and was eventually adopted in consumer electronics. It allowed for the miniaturization of devices such as personal computers, mobile phones, and other consumer gadgets.

Facilitating Moore's Law—By enabling higher transistor densities and more efficient heat dissipation, C4 played a role in sustaining Moore's Law, which predicts the doubling of transistor counts approximately every two years.

Take Away

C4 technology, developed in the 1960s by IBM, was a groundbreaking innovation in semiconductor packaging. By enabling direct, reliable connections between the semiconductor die and the substrate, it allowed for full utilization of the advantages of higher transistor densities, produced better performance, and yielded increased reliability in electronic devices. One drawback of this technology consisted in the much higher cost as compared with the dominant packaging technologies of the time like wire bonding.

For this reason, C4 was initially used in IBM’s mainframe computers and high-performance processors but C4 technology quickly found applications in military and aerospace where the benefits outweighed the increased packaging cost but slowly eventually C4 became a standard in the broader semiconductor industry. The development and early applications of C4 technology marked a significant milestone in the evolution of electronic packaging, influencing the design and manufacturing of integrated circuits for decades to come. Again, the main drawback of C4 technology consisted in unusually high cost in the world of packaging.

As time went by this cost became more affordable as electronics products reached much higher selling prices (i.e., thousands of dollars) since it represented an affordable amount in the bill of materials (BOM) than in the past when it was only justified in cases of large expensive mainframe computers.

It took another 30 years from the first introduction before a new large electronics market driven by Data Centers could eventually afford this expensive but extremely useful packaging technology.

Flip Chip Packaging Fundamentals

Flip chip packaging is an advanced semiconductor packaging technology that offers several advantages in terms of performance and miniaturization. It involves flipping the semiconductor chip upside down and directly bonding it to the package substrate or to another chip using solder bumps. Figure 4 illustrates a typical flip chip. Below is an overview of flip chip technology. Table 12 lists the advantages and disadvantages of flip chip packaging. Appendix D details flip chip manufacturing.

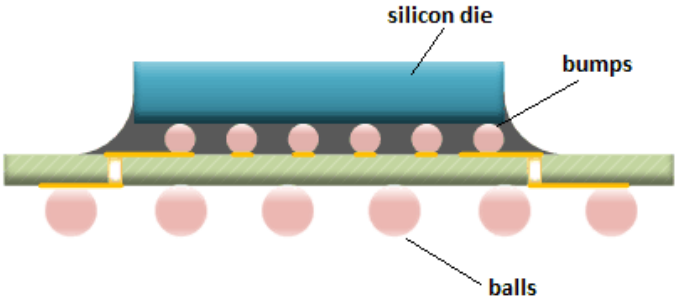


Figure 4. Example of Flip Chip packaging on substrate

Table 12. Advantages and Disadvantages of Flip Chip Packaging

Advantage	Description
Enhanced Electrical Performance	• **Shorter Electrical Paths**: Flip chip technology reduces the distance between the chip and the package substrate, which minimizes signal delay and improves performance. • **Reduced Inductance and Capacitance**: The direct connection between the chip and substrate lowers inductance and capacitance, which enhances high-speed signal integrity.
Improved Thermal Management	• **Better Heat Dissipation**: Flip chip packaging often allows for better thermal conductivity and heat dissipation due to the direct connection between the chip and the substrate.
Higher I/O Density	• **Increased Interconnection**: The use of solder bumps enables a higher density of electrical connections compared to traditional wire-bonding methods, which supports more complex and higher-performance devices.
Smaller Form Factor	• **Compact Design**: Flip chip technology supports more compact and lightweight package designs, making it suitable for applications where space is at a premium.
Improved Reliability	• **Reduced Mechanical Stress**: The direct bonding of the chip to the substrate reduces the mechanical stress associated with wire bonds, which can enhance the reliability of the package.

Applications of Flip Chip Packaging

High-Performance Computing—Used in processors and memory modules where high-speed data transfer and efficient thermal management are critical.

Consumer Electronics—Common in smartphones, tablets, and other compact electronic devices where space and performance are key considerations.

Graphics Processing Units—Flip chip technology is used in GPUs to handle high-speed data processing and cooling requirements.

Networking and Communication Devices—Employed in high-speed networking equipment where performance and density are important.

Automotive Electronics—Used in automotive applications where reliability and durability are crucial.

Flip Chip Packaging Challenges and Considerations

Cost—Flip chip packaging can be more expensive than traditional packaging methods due to the complexity of the process and materials used.

Manufacturing Complexity—The flip chip process requires precise alignment and reflow processes, which can be challenging to control.

Thermal Management—While flip chip can improve thermal management, it still requires effective heat dissipation solutions to handle high-performance applications.

Design and Reliability—Designing for flip chip packaging requires careful consideration of thermal and mechanical factors to ensure long-term reliability.

Overall, flip chip packaging technology is a powerful solution for modern electronic IC products that can finally afford the cost of FC to satisfy demand for high performance, reliability, and miniaturization.

Brief History of Evolutions of Bump Composition in Flip Chip

The evolution of bump composition in C4 (Controlled Collapse Chip Connection) flip-chip technology has been driven by the need to improve performance, reliability, and compatibility with different semiconductor materials and packaging processes. Here's an overview of how bump compositions have evolved over time:

Early C4 Bumps (1960s–1980s)

- Material—The original C4 bumps developed by IBM in the 1960s primarily used high-lead solder (95Pb/5Sn).
- Characteristics—High-lead solders were favored due to their high melting point (around 300°C) and robust mechanical properties, which were suitable for the thermal expansion mismatches between silicon chips and ceramic substrates.
- Challenges—Over time, concerns arose regarding the environmental and health impacts of lead-based solders, prompting the industry to seek alternatives.

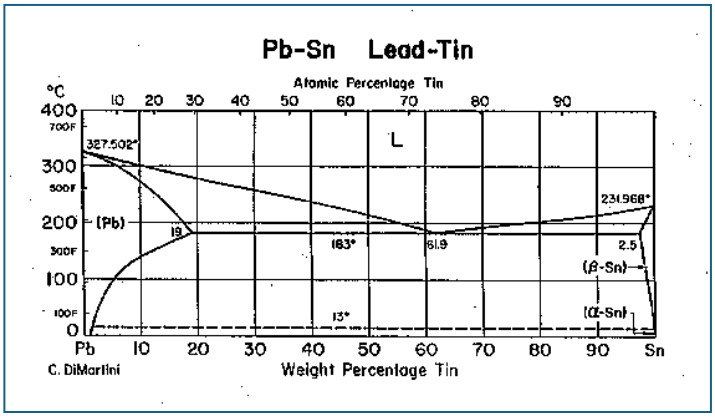


Figure 5. Phase diagram of Pb-Sn showing composition vs. collapse temperature

Transition to Lead-Free Solder (1990s–2000s)

- Material—As regulations like the European Union's RoHS (Restriction of Hazardous Substances) directive came into force, the industry began transitioning to lead-free solders. Common alternatives included Sn/Ag/Cu (SAC) alloys, with compositions like SAC305 (96.5Sn/3Ag/0.5Cu).
- Characteristics—Lead-free solders have a lower melting point (around 217-220°C) compared to high-lead solders and provide better compatibility with organic substrates used in modern packaging.
- Challenges—The transition to lead-free materials required adjustments in manufacturing processes, as the mechanical properties and reliability differed from those of lead-based solders.

Advanced Bump Compositions (2010s–Present)

- Material—In recent years, there has been an increased focus on improving the reliability and performance of bumps in advanced packaging. Low-Ag SAC alloys (e.g., SAC105), anisotropic conductive films (ACF), and copper (Cu) pillars have gained popularity.

- Copper Pillars—A significant innovation, Cu pillars are used with a thin layer of solder on top, allowing for finer pitch interconnects and improved electrical performance. Cu pillars are particularly advantageous in high-density applications like 2.5D and 3D ICs.
- High-Performance Solders—For specific applications requiring enhanced thermal and mechanical performance, high-reliability alloys with small additions of elements like Bi, Sb, or Ni have been introduced.

4. Future Trends (2020s–2030s)

- Material—The industry is moving towards even more advanced bump materials and structures. Hybrid bonding and solid-state diffusion bonding are emerging as potential replacements for traditional solder bumps in 3D IC packaging, promising lower resistance and higher reliability.
- Characteristics—These new methods could lead to reduced bump sizes, enabling further miniaturization and higher interconnect densities, crucial for next-generation semiconductor technologies.

Summary of Evolution

- 1960s–1980s—High-lead solder (95Pb/5Sn)
- 1990s–2000s—Transition to lead-free solder (SAC alloys)
- 2010s–Present—Introduction of low-Ag SAC, Cu pillars, and high-performance solders
- 2020s–2030s—Exploration of hybrid bonding and solid-state diffusion bonding

This evolution reflects the industry's ongoing efforts to meet the demands of smaller, faster, and more reliable semiconductor devices while addressing environmental concerns and regulatory requirements.

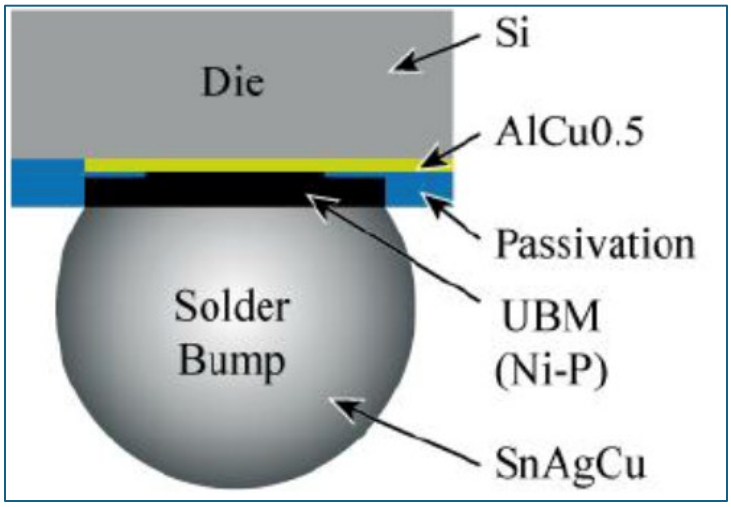


Figure 8. Modern bump composition

Table 13. Advantages of Flip Chip Process Technology

Process	Description
Improved Electrical Performance	• Shorter interconnect lengths and reduced inductance and capacitance contribute to better signal integrity and higher performance.
Enhanced Thermal Management	• Direct contact with the substrate allows for more efficient heat dissipation compared to traditional wire bonding.
Compact Package Size	• Flip chip technology allows for higher I/O density and smaller package sizes, which is beneficial for high-performance and space-constrained applications.

Conclusion

The flip chip manufacturing process involves preparing the wafer with bumps, flipping the chip to align it with the substrate, bonding it using solder bumps, and applying and curing underfill material. The process is followed by dicing, encapsulation, and final testing to ensure functionality and reliability. Flip chip technology offers advantages in electrical performance, thermal management, and package size, making it suitable for advanced and high-performance semiconductor applications.

Flip Chip Materials

In flip chip packaging, several materials are utilized to ensure reliable electrical connections, thermal management, and overall performance. Here's a detailed look at the key materials involved in flip chip technology in Figure 9 and Table 14.

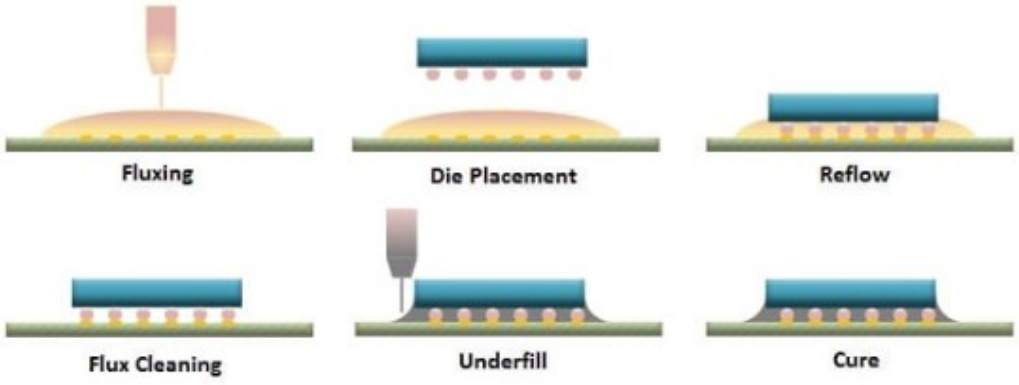


Figure 9. Step-by-step schematic overview of the fabrication of the flip chip process

Table 14. Flip chip process elements materials and purpose

Flip Chip Process Element	Material Composition and Purpose
Solder Bumps	• Material: Originally composed of an alloy of tin (Sn) and lead (Pb) or, more commonly in recent years, lead-free alloys such as tin-silver-copper (SAC) alloys. • Purpose: Serve as the electrical and mechanical connection between the chip and the package substrate or another chip. The bumps are reflowed to create a solid connection.
Die Attach Adhesives	• Material: Includes a variety of materials such as epoxy resins, silver-filled epoxies, or other thermally conductive adhesives. • Purpose: Bond the semiconductor die to the package substrate or to an intermediate layer. They are chosen for their thermal and mechanical properties.
Underfill Materials	• Material: Typically epoxy-based resins, sometimes filled with silica or other fillers to improve mechanical properties. • Purpose: Applied between the chip and the substrate after solder bump reflow to fill the gaps and improve the mechanical robustness and thermal conductivity of the flip chip assembly.
Passivation Layers	• Material: Often made from silicon dioxide (SiO_2) or silicon nitride (Si_3N_4). • Purpose: Protects the semiconductor surface from contaminants and physical damage. These layers are applied on the die before the flip chip bonding process.

Flip Chip Process Element	Material Composition and Purpose
Redistribution Layers (RDLs)	• Material: Typically consists of multiple metal layers, such as copper (Cu), deposited on a dielectric layer. • Purpose: Used to route electrical signals from the die to the external package connections. RDLs are crucial for managing complex interconnections and improving signal integrity.
Thermal Interface Materials (TIMs)	• Material: Includes thermal pads, thermal grease, or phase-change materials. • Purpose: Improve the thermal conductivity between the die and the heat sink or package to manage heat dissipation.
Substrate Materials	• Material: Commonly used substrates include organic laminates (such as FR-4 or BT resin-based materials) or ceramic materials. • Purpose: Provides mechanical support and electrical connections for the flip chip. The choice of substrate impacts the overall performance, size, and cost of the package.
Encapsulation Materials	• Material: Typically composed of epoxy resins or other protective polymers. • Purpose: Protects the entire flip chip assembly from physical damage, moisture, and other environmental factors. Encapsulation also helps in providing additional thermal management.
Solder Mask	• Material: Usually made from epoxy-based or UV-curable resins. • Purpose: Applied to the PCB or substrate to protect the copper traces and prevent solder from bridging unwanted areas.
Dielectric Materials	• Material: Includes materials like polyimide or other high-temperature-resistant polymers. • Purpose: Used as insulating layers between different metal layers or between the die and the substrate to prevent electrical shorts and improve performance.

Key Considerations for Materials

- **Thermal Management—**Effective heat dissipation is crucial, so materials with good thermal conductivity are selected.
- **Mechanical Strength—**The materials must withstand mechanical stresses during handling and operation.
- **Electrical Performance—**The materials need to support high-speed signal transmission with minimal interference.
- **Reliability—**Long-term reliability of the flip chip package is dependent on the durability and stability of the materials used.

Overall, the choice of materials in flip chip packaging is critical to ensuring high performance, reliability, and durability of the final semiconductor device. Each material is selected based on its specific properties and its role in the overall flip chip assembly.

Flip Chip Bump Spacing (Pitch) Requirements

The distance between bumps, often referred to as bump pitch, is a critical parameter in packaging design, particularly in microelectronic and semiconductor applications. The bump pitch is the center-to-center distance between adjacent bumps or solder balls on a package. The specific distance between bumps can vary widely depending on the application, packaging technology, and design requirements.

Factors Influencing Bump Pitch

- **Application Requirements**—Different applications have different requirements for bump pitch. For example, high-density memory devices or processors may require smaller bump pitches to fit more connections in a limited space, whereas less dense applications might use larger pitches.

- **Packaging Type**—The type of packaging influences bump pitch. Common packaging types include:
  - **Flip-Chip Packages**—Typically use smaller bump pitches to accommodate high-density interconnections. Standard pitches can range from 0.4 mm to 0.8 mm, with advanced applications pushing below 0.4 mm.
  - **Ball Grid Array (BGA)**—In BGAs, bump pitches can vary from 0.8 mm to 1.0 mm or larger, depending on the package size and component density.
- **Manufacturing Capabilities**—Advances in manufacturing technologies allow for finer bump pitches. Techniques such as precision photolithography, advanced solder paste printing, and micro-manipulation contribute to achieving smaller pitches with high accuracy.
- **Thermal and Electrical Performance**—The spacing between bumps can affect the thermal and electrical performance of the package. Smaller bump pitches can increase thermal density and improve electrical performance but may require careful management of heat dissipation and signal integrity.
- **Cost and Complexity**—Smaller bump pitches can increase manufacturing complexity and cost. Balancing performance with cost considerations is crucial. Technologies for fine bump pitches are more expensive due to the precision required in both manufacturing and inspection.

Trends and Examples

- **Fine Pitch Technology**—The trend towards miniaturization in electronics drives the use of finer bump pitches. For instance, in advanced flip-chip packages, bump pitches as small as 0.3 mm or even 0.2 mm are becoming more common.
- **3D Packaging**—In 3D packaging technologies, such as through-silicon vias and stacked die packages, bump pitches are often very fine to accommodate the dense interconnects between multiple layers. These can be as small as 0.1 mm in cutting-edge designs.
- **Advanced Manufacturing**—Innovations in manufacturing processes, such as ultra-fine photolithography and precision placement techniques, enable the production of smaller bump pitches with high reliability.
- **Emerging Applications**—New applications in areas like high-performance computing, mobile devices, and automotive electronics are driving the need for smaller bump pitches to accommodate increasing functionality in compact spaces.

In summary, bump pitch is a crucial design parameter in packaging that balances density, performance, and cost. The trends in bump pitch are driven by the need for smaller, more efficient, and higher-performance electronic devices, supported by advancements in manufacturing technology and application requirements.

Recent Trends in Flip Chip Bump Spacing

In recent years, there have been several notable trends in packaging technology that relate to bump spacing, which can impact both the functionality and aesthetics of packaging design. Here are a few key trends:

- **Minimalist Design**—Packaging is increasingly moving towards minimalist aesthetics. This often means using fewer design elements, including bump or texture patterns, to create a sleek and modern look. Minimalist packaging focuses on clean lines and simple textures.
- **Sustainability**—As brands strive to reduce their environmental footprint, there's a growing emphasis on using eco-friendly materials and processes. This can affect bump spacing and design as companies seek to balance sustainability with functionality. For instance, lighter embossing or debossing might be used to save on materials and energy.
- **Functional Textures**—Bump spacing and textures are being used to enhance functionality. For example, raised or textured areas can provide better grip on packaging, which is particularly important for products that might be used in wet or slippery conditions.
- **Interactive and Tactile Experiences**—There's a trend towards creating packaging that engages multiple senses. Bump spacing and tactile elements can make the packaging experience more interactive. For example, products might feature tactile patterns that encourage consumers to touch and feel, enhancing their connection with the product.

- **Personalization**—With advances in printing and manufacturing technology, personalized packaging has become more popular. This can include custom bump patterns or textures that cater to individual preferences or special editions of products.
- **Smart Packaging**—The integration of technology into packaging is growing. Bumps and textures can be used in conjunction with QR codes, NFC tags, or other smart features to enhance the user experience or provide additional information.
- **Aesthetic Variety**—While some trends emphasize minimalism, others explore a range of textures and bump spacings to create visually interesting and unique packaging. This can include complex patterns or irregular textures that make the packaging stand out on the shelf.

These trends reflect broader shifts in consumer preferences and technological advancements in the packaging industry. They indicate a move towards designs that are not only visually appealing but also functional and aligned with current values, such as sustainability and personalization.

Historical Trends in Bump Spacing

The trend in silicon packaging bump spacing has been characterized by a continuous decrease in bump pitch as the semiconductor industry pushes towards higher density, increased performance, and reduced power consumption in electronic devices. This trend is driven by the need for more I/O connections, smaller package sizes, and improved electrical and thermal performance.

- **2010s**—Bump pitches in advanced packaging technologies such as flip-chip and wafer-level packages were typically around 150-200 micrometers (µm).
- **2020**—By the early 2020s, bump pitches had reduced to approximately 40-100 µm in cutting-edge applications, particularly in high-performance computing, mobile processors, and other advanced semiconductor packages.
- **2025 Projection**—Industry projections have shown bump pitches reaching as low as 30-50 µm for leading-edge silicon packages, driven by advancements in lithography, bumping technology, and finer pitch redistribution layers (RDLs).

Key Drivers for Continued Reduction

- **Increased I/O Density**—As chip designs continue to demand more I/O connections to support advanced functionalities, the number of bumps per unit area must increase, necessitating tighter bump pitches.
- **Advanced Packaging Techniques**—Technologies like through-silicon vias, 2.5D/3D integration, and Fan-Out Wafer-Level Packaging (FOWLP) push the limits of bump pitch reduction by allowing more
- compact and efficient interconnects.

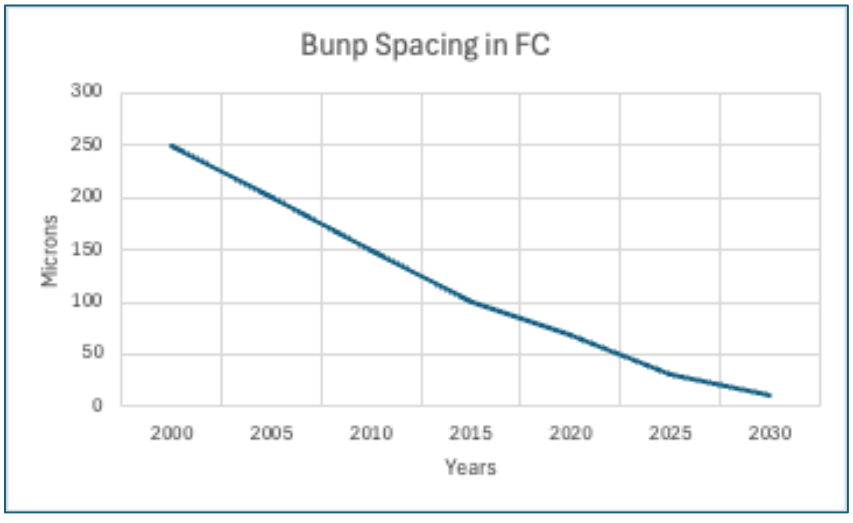


Figure 10. Historical trend of bump spacing and forecast

Process Node Shrinks—As semiconductor process nodes continue to shrink (e.g., 5nm, 3nm, and beyond), the bump pitch must also scale down to match the finer features of the ICs, enabling denser and more power-efficient packaging.

Extrapolation to 2030

- **Bump Pitch Trends**—Based on the ongoing trends and technological advancements, it is reasonable to expect that bump pitches could reach 10-20 µm or even smaller by 2030. This extrapolation considers the continued development of ultra-fine lithography techniques, advanced deposition methods, and enhanced materials that can support such small geometries.
- **Advanced Interconnect Solutions**—To support these smaller pitches, new interconnect solutions such as hybrid bonding (direct Cu-Cu bonding) and advanced polymer dielectrics may become more prevalent, enabling further reductions in bump spacing while maintaining or even improving reliability and performance.

Challenges and Considerations

- **Manufacturing Complexity**—As bump pitches continue to shrink, the manufacturing process becomes increasingly complex and requires more precise control over materials and processes. This includes challenges in lithography, bump deposition, and alignment.
- **Yield Management**—The smaller the bump pitch, the more susceptible the process is to defects and misalignments, which can significantly impact yield. Advanced inspection and metrology tools will be essential to maintain acceptable yields.
- **Thermal Management**—As bump spacing decreases, thermal management becomes more critical due to the increased density of power dissipation. Innovations in thermal interface materials (TIMs) and package design will be necessary to handle these challenges.

Conclusion

The trend in bump spacing (pitch) in FC packaging is expected to continue its downward trajectory, reaching pitches as small as 10-20 µm by 2030. This reduction is driven by the need for higher I/O density, advanced packaging techniques, and ongoing semiconductor scaling. However, achieving these smaller bump pitches will require overcoming significant challenges in manufacturing, yield, and thermal management. The industry will need to continue innovating in both materials and processes to realize the potential of ultra-fine bump pitches in the next decade.

Miniaturization of Bumps for Flip Chip

In micron packaging, which often involves packaging for electronic components, precision instruments, or high-tech products, bump spacing trends are closely related to technological advancements and specific application requirements. Here are some key trends in bump spacing for micron-level packaging:

Table 15. Bump Spacing Trends for Micro-level Packaging

Trend	Description
Increased Miniaturization	• As electronics become more compact, the need for finer bump spacing in packaging grows. This trend is driven by the demand for smaller, more powerful devices with higher performance and increased functionality. Bump spacings are becoming smaller to accommodate denser component layouts and reduce the overall footprint of the packaging.
Advanced Materials	• The use of advanced materials, such as high-density interconnect (HDI) substrates and new types of solder materials, is influencing bump spacing. These materials can support finer bump spacings while maintaining reliability and performance. Innovations in materials help achieve more precise and consistent bump placements.

Trend	Description
Improved Manufacturing Techniques	• Advances in manufacturing processes, such as photolithography and precision stencil printing, allow for tighter control over bump spacing. Enhanced accuracy in these processes ensures that smaller and more closely spaced bumps can be reliably produced, which is crucial for high-density applications.
Thermal Management	• In high-performance electronics, managing heat dissipation is crucial. Trends in bump spacing are sometimes influenced by the need to optimize thermal performance. Engineers may adjust bump spacing to enhance heat dissipation or integrate thermal management solutions within the packaging design.
High-Reliability Requirements	• In applications where reliability is critical, such as aerospace or medical devices, bump spacing must be carefully controlled to ensure durability and performance under various conditions. This trend involves tighter specifications and more stringent quality control measures to meet high-reliability standards.
Cost and Performance Trade-offs	• As bump spacing decreases, the cost of manufacturing often increases due to the need for more precise equipment and materials. There is a continuous balancing act between achieving the desired performance and managing costs. Trends in bump spacing often reflect these trade-offs and the drive for cost-effective yet high-performance solutions.
Integration of New Technologies	• Emerging technologies, such as advanced flip-chip bonding and 3D packaging, are pushing the boundaries of bump spacing. These technologies require precise bump placement to ensure proper electrical connections and performance in increasingly complex and miniaturized devices.

These trends reflect the ongoing evolution in micron-level packaging, driven by the need for smaller, more efficient, and higher-performance electronic devices. As technology continues to advance, bump spacing will likely keep evolving to meet the demands of new applications and innovations.

Fifth ETP Summary—Flip Chip evolution

In the 60s, as transistor counts increased and circuits became more complex, the limitations of traditional wire bonding became apparent. Wire bonds were prone to damage, could only connect to the edges of the die, and had significant limitations in terms of density and performance, only a limited number of connections could be made from the package to the outside world. The core idea behind C4 technology was to "flip" the semiconductor die upside down and directly bond it to the substrate using small solder bumps. This allowed for connections to be made across the entire surface of the die rather than just the edges. However, the technology was very expensive as compared to the packaging cost that consumer products could afford at the time. In fact, the first commercial use of C4 technology was the System/360, introduced in 1964, and later in the System/370 by IBM. It would then take another 50 years before the technology became relevant and affordable for high-performance computing.

6. Single Die versus Multi-Chips Assembly Module

When choosing between single-die (SoC) and multi-chip (SiP) configurations, several trade-offs must be considered, including performance, cost, complexity, and thermal management. The following is a detailed comparison of the two approaches:

Single Die Configuration

Table 16. Advantages and Disadvantages of Single Die Configuration

Configuration	Advantages	Disadvantages
Single Die	• **Simplicity:** – **Design and Manufacturing:** Single-die packages are generally simpler to design and manufacture. They do not require complex inter-die communication or alignment processes. – **Integration:** Since all functions are integrated into one chip, there are fewer concerns about inter-die compatibility or alignment issues. • **Performance:** – **Lower Latency:** Direct connections within a single die result in lower latency and higher speed due to shorter signal paths. – **Consistency:** Performance is more consistent as all components operate within the same silicon die, leading to uniformity in thermal and electrical characteristics. • **Cost**: – **Lower Packaging Cost:** Single-die packages generally have lower packaging costs since they avoid the complexities associated with multi-chip packaging. • **Power Efficiency:** – **Reduced Power Consumption:** Fewer interconnects and simpler design can lead to better power efficiency, as there are fewer power losses associated with long signal paths.	• Design Limitations: – Size Constraints: The size of the single die is limited by current manufacturing technologies. This can restrict the complexity or functionality that can be integrated into one die. – Yield Issues: Larger die sizes can lead to lower manufacturing yields, as defects are more likely to impact larger chips. • Flexibility: – Limited Upgradability: Upgrading or modifying specific functions within a single die is challenging. Any changes typically require redesigning the entire chip.

Multi-Chip Configuration

Table 17. Advantages and Disadvantages of Multi-Chip Configuration

Configuration	Advantages	Disadvantages
Multi-Chip	• **Scalability:** – **Increased Functionality:** Multi-chip packages (MCPs) allow for combining different types of chips (logic, memory, analog) in one package, enabling more complex and feature-rich designs. – **Flexibility:** Different chips can be developed and upgraded independently, providing greater flexibility in design and development. • **Cost Efficiency in Production:** – **Yield Improvement:** Smaller individual chips typically have higher yields than larger single dies. This can make multi-chip configurations more cost-effective in high-volume production. • **Customization:** – **Tailored Solutions:** Allows for creating custom solutions by combining specific chiplets or components tailored to particular applications. • **Thermal Management:** – **Better Heat Distribution:** Multi-chip configurations can be designed to better manage heat, with different chips handling different functions and thus distributing thermal loads more effectively.	• **Complexity:** – **Design Complexity:** Multi-chip packages are more complex to design and manufacture due to the need for precise alignment, interconnects, and thermal management. – **Interconnects:** Requires effective management of inter-chip communication, which can introduce latency and complexity. • **Cost:** – **Higher Packaging Costs:** Multi-chip packages may have higher packaging and assembly costs due to the need for advanced packaging technologies and processes. • **Power Consumption:** – **Potential Power Overhead:** More complex interconnects and multiple chips can lead to higher power consumption and potential inefficiencies in signal routing.

Choosing Between Single Die and Multi-Chip Configurations

The decision between single-die and multi-chip configurations depends on various factors including:

- **Application Requirements**—For high-performance applications needing maximum speed and efficiency, a single die may be preferred. For applications requiring a high degree of functionality and flexibility, a multi-chip approach may be more suitable.
- **Cost Considerations**—Evaluate both manufacturing and packaging costs. Multi-chip solutions might be more cost-effective in some scenarios due to improved yields.
- **Design Constraints**—Consider the size limitations of single-die solutions and the complexity of multi-chip designs.
- **Thermal Management**—Assess the thermal requirements and how each approach handles heat dissipation.

In summary, single-die solutions offer simplicity and efficiency, but are limited by size and design constraints. Multi-chip configurations provide flexibility and scalability but come with increased complexity and potential cost considerations. The best choice depends on the specific needs and constraints of the application.

Overview of System in Package

Analysis of Tradeoffs of SiP Technical and Cost Benefits

High-Level Integration

- **Homogeneous and Heterogeneous Integration**—Silicon SiP allows homogeneous integration of diverse products made with semiconductor technologies—such as processors, memory, RF components, and sensors—into a single package. Heterogeneous integration enables the creation of compact, multifunctional devices that combine silicon and non-silicon technologies maximizing the cooperative strengths of different components.
- **Miniaturization**—By integrating multiple ICs into a single package, silicon SiP significantly reduces the overall size of the system. This is crucial for applications like mobile devices, wearables, and IoT devices, where space is limited.

Performance Enhancement

- **Improved Signal Integrity**—Silicon SiP minimizes the length of interconnections between components by integrating them closely within the package. This reduces signal loss, crosstalk, and latency, leading to enhanced performance, especially in high-speed applications.
- **Optimized Power Consumption**—The integration of multiple components within a single package enables more efficient power distribution and management, leading to lower overall power consumption. This is particularly beneficial in battery-powered devices.

Flexibility and Customization

- **Design Flexibility**—Silicon SiP technology offers significant design flexibility, allowing manufacturers to customize the package to meet specific application requirements. This includes the ability to integrate a mix of analog, digital, RF, and MEMS (microelectromechanical systems) components.
- **Tailored Solutions**—SiP can be tailored to address specific market needs, such as high-performance computing, advanced telecommunications, or automotive applications. This customization enables manufacturers to offer differentiated products that cater to particular segments.

Cost Efficiency

- **Reduced Bill of Materials (BOM)**—By integrating multiple functions into a single package, silicon SiP reduces the need for discrete components and complex PCB (printed circuit board) designs. This leads to a lower bill of materials and simplified assembly processes, reducing overall manufacturing costs.
- **Economies of Scale**—Silicon SiP allows for the reuse of proven semiconductor processes and IP blocks, enabling cost-effective production at scale. This is particularly advantageous for high-volume markets such as consumer electronics.

Advanced Packaging Technologies

- **Through-Silicon Vias**—Silicon SiP often employs TSVs to connect multiple layers of silicon within the package, enabling 3D integration. This advanced packaging technique enhances performance by providing high-density interconnections and reducing signal delays.
- **Fan-Out Wafer-Level Packaging (FOWLP)**—FOWLP is another advanced packaging technology used in silicon SiP, allowing for the creation of very thin and compact packages with excellent thermal and electrical performance. This is especially useful in mobile and IoT applications.

Market Demands and Applications

- **5G and Telecommunications**—The demand for 5G and other advanced telecommunications technologies drives the need for silicon SiP, which can integrate RF components, antennas, and baseband processors into a single package. This integration is essential for the compact and efficient design of 5G devices.
- **Automotive and Industrial Applications**—In automotive and industrial sectors, silicon SiP is crucial for creating robust, high-performance systems that can operate in harsh environments. SiP enables the integration of sensors, microcontrollers, and communication modules in compact, reliable packages.

- **Consumer Electronics**—The push for thinner, more powerful consumer electronics—such as smartphones, tablets, and wearables—drives the adoption of silicon SiP. The technology allows for high-performance, multifunctional devices that are both compact and energy-efficient.

Reduced Time-to-Market

- **Rapid Prototyping and Testing**—Silicon SiP technology enables quicker prototyping and testing compared to traditional multi-chip solutions. This accelerates the product development cycle, allowing manufacturers to bring new products to market faster.
- **Modular Approach**—SiP's modular nature allows for the reuse of existing components and designs, further reducing development time and enabling faster adaptation to changing market demands.

Cost Implications of Silicon SiP Technology

While silicon SiP offers numerous advantages, its development and implementation also come with specific cost considerations, as follows:

High Initial Development Costs

- **Complex Design and Integration**—Designing a silicon SiP package involves significant R&D investment, particularly in the areas of heterogeneous integration and advanced packaging technologies like TSVs and FOWLP. The complexity of integrating multiple components into a single package increases design costs.
- **Advanced Manufacturing Processes**—Manufacturing silicon SiP packages requires advanced processes that are more costly than traditional packaging methods. These processes include precision lithography, wafer thinning, and TSV creation, which add to the overall cost.

Manufacturing and Testing Costs

- **Yield Management**—The integration of multiple components within a silicon SiP package can lead to yield challenges, as the failure of any single component may affect the entire package. This necessitates stringent quality control and testing procedures, which can increase manufacturing costs.
- **Advanced Packaging Equipment**—The need for specialized equipment to manufacture and test silicon SiP packages, such as TSV etching tools and wafer-level test systems, contributes to higher capital expenditures.

Economies of Scale

- **Volume Production**—To achieve cost-effectiveness, silicon SiP production must be scaled up to high volumes. This requires substantial initial investment but can lead to lower per-unit costs as production ramps up.
- **Market Demand**—The success of silicon SiP in achieving economies of scale depends on strong market demand. Manufacturers must carefully assess market trends and consumer needs to ensure sufficient demand for their SiP products.

Risk of Technological Obsolescence

- **Rapid Technological Changes**—The fast pace of technological advancements in the semiconductor industry means that silicon SiP designs can quickly become obsolete. Continuous investment in R&D is required to keep up with new developments, adding to long-term costs.
- **Market Shifts**—Market dynamics, such as shifts in consumer preferences or new technological standards, can impact the demand for specific silicon SiP solutions. Companies must be agile in adapting their SiP offerings to meet changing market needs.

Conclusion

The motivations and drivers for silicon SiP technology are centered around the need for high integration, improved performance, flexibility, and cost efficiency in modern electronic devices. Silicon SiP enables the creation of compact, powerful, and multifunctional systems that meet the demands of advanced applications, particularly in telecommunications, automotive, and consumer electronics. However, the adoption of silicon SiP also involves significant cost implications, including high development and manufacturing expenses, the need for advanced packaging technologies, and the challenge of achieving economies of scale. Despite these costs, the benefits of silicon

SiP—such as reduced size, enhanced performance, and faster time-to-market—make it a critical technology in the evolution of electronics.

Multi-Chip Module Assembly

Multichip module assembly involves integrating multiple semiconductor chips into a single package or module. This process is used to enhance functionality, increase performance, and reduce the size of electronic devices. Refer to Appendix F for MCM assembly process details.

Additional Considerations for MCMs

- **Thermal Management—**Implement effective thermal management solutions to address heat generated by multiple chips, such as heat sinks or thermal vias.
- **Electrical Performance—**Ensure that the electrical interconnections are optimized to minimize signal degradation and maintain high performance.
- **Reliability—**Focus on reliability and durability to ensure that the multichip assembly performs reliably over its expected lifetime.

In summary, MCM technology is driven by cost and performance. By assembling multiple dice on a single substate reduced packaging costs. In addition, it greatly benefits reduction in signal propagations delays as the dice spacing is substantially reduced. The specific details and techniques in the multichip assembly process can vary based on the type of chips being used, the complexity of the design, and the intended application.

MCM Interconnections to Board

In MCM assemblies, the substrate is a critical component that supports and interconnects multiple integrated circuits within a single package. The type and preparation of the substrate play a significant role in the performance, reliability, and functionality of the MCM. Below is a detailed look at substrate types and preparation processes for multi-chip assemblies:

Substrate Types for Multi-Chip Assembly

1. **Printed Circuit Board Substrates**
    - **Material**—Typically made from FR-4 (fiber-reinforced epoxy laminate) or other composite materials. For high-performance applications, materials with better thermal and electrical properties, such as polyimide or ceramic-filled epoxy, may be used.
    - **Use**—Common in lower complexity MCMs or in cases where cost-effectiveness is a priority. PCBs are often used in consumer electronics and less demanding applications.
2. **Ceramic Substrates**
    - **Material**—Made from materials like alumina (Al_2O_3), aluminum nitride (AlN), or beryllium oxide (BeO). These materials offer high thermal conductivity and excellent electrical insulation.
    - **Use**—Preferred for high-power or high-frequency applications where superior thermal management and electrical performance are required. Used in applications such as telecommunications and aerospace.
3. **Metal-Organic Frameworks (MOFs)**
    - **Material**—These are less common but include materials like copper or stainless steel with a combination of organic and inorganic components.
    - **Use**—Useful for applications requiring robust mechanical strength and high thermal conductivity.
4. **Glass Substrates**
    - **Material**—Made from specialized glass with good electrical insulating properties.
    - **Use**—Typically used in high-frequency applications due to their low dielectric loss and good thermal stability.

5. **Organic Laminates**
    - **Material**—Advanced organic laminates, such as BT (bismaleimide-triazine) resin, are used for high-density interconnects.
    - **Use**—Suitable for high-density and high-performance MCMs, often used in advanced electronics and communications.

Details for Substrate Preparation Process for Multi-Chip Assembly is presented in Appendix G.

Conclusion

The substrate for multi-chip assemblies can vary based on the requirements of the application, ranging from standard PCBs to advanced ceramic or glass substrates. The preparation of the substrate involves several critical steps, including design, layer formation, via creation, solder masking, and surface finishing.

Proper substrate preparation is crucial for ensuring the reliability and performance of the multi-chip module, making it suitable for complex and high-performance applications.

Summary of Typical SiP Technologies

The development and adoption of silicon-based SiP technology are driven by various factors that focus on enhancing the integration of multiple components, improving performance, reducing size, and enabling advanced functionalities. Silicon SiP technology is increasingly important in applications where traditional single-chip solutions are insufficient, particularly in advanced computing, telecommunications, consumer electronics, and automotive industries. Semiconductor packaging is evolving rapidly due to the increasing demands for smaller, faster, and more energy-efficient electronic devices. The following are important key trends expected to shape the future of semiconductor packaging:

Advanced Packaging Technologies

- **2.5D and 3D Packaging**—These technologies enable stacking multiple chips on top of each other or side by side, leading to better performance and lower power consumption. 2.5D involves placing chips side by side on an interposer, while 3D involves stacking them vertically.
- **Chiplet Architecture**—Instead of designing a single large monolithic chip, chiplets allow smaller, function-specific chips to be combined, offering flexibility and cost advantages.
- **Fan-Out Wafer-Level Packaging (FOWLP)**—This packaging technique offers a higher level of miniaturization and better thermal performance, making it suitable for high-performance computing and mobile devices.
- **Homogeneous Integration**
    - Integration of silicon-based technologies—Different dice produced with variations of the basic silicon technology like memory, logic, application specific IC (ASIC), etc. are integrated by placing and interconnecting them in a single package
- **Heterogeneous Integration**
    - Integration of Diverse Technologies—Combining different types of chips (e.g., analog, digital, photonic) into a single package allows for more complex and capable systems. This trend is crucial for applications like AI, IoT, and 5G.

System considerations-in SiP

Miniaturization of Entire Systems—SiP packages multiple components (e.g., processors, memory, sensors) into a single module, enabling compact devices with enhanced functionality. This trend is particularly important for wearable technology and IoT devices.

Thermal Management Innovations

- **Advanced Cooling Solutions**—As chips become more powerful, managing heat dissipation is critical. Techniques like integrated liquid cooling, improved thermal interface materials, and innovative packaging designs are gaining importance.

Materials Advancements

- **Use of New Materials**—The shift to new materials like silicon carbide (SiC) and gallium nitride (GaN) is enabling higher efficiency and better performance, particularly in power electronics and RF applications.
- **Bio-compatible Materials**—For medical devices and wearables, there's a growing demand for packaging materials that are safe for long-term contact with the human body.

Reliability and Testing

- **Improved Testing Techniques**—As packaging becomes more complex, so does the need for advanced testing methods to ensure reliability, especially for critical applications in automotive, aerospace, and healthcare.
- **Enhanced Reliability**—Packaging designs are increasingly focusing on improving reliability under harsh conditions, such as high temperatures and mechanical stress, which is critical for automotive and industrial applications.

Sustainability and Cost Efficiency

- **Eco-friendly Packaging**—The semiconductor industry is pushing towards more sustainable practices, including the use of recyclable materials, reduction of hazardous substances, and energy-efficient manufacturing processes.
- **Cost-Effective Manufacturing**—The ongoing trend towards cost reduction will drive innovations in packaging processes that can reduce overall manufacturing costs without compromising performance.

AI and Machine Learning in Packaging Design

- **Optimization through AI**—Artificial intelligence and machine learning are increasingly being used to optimize packaging design and manufacturing processes, leading to better performance and reduced time-to-market.

Photonics Integration

- **Photonic Packaging**—With the growing importance of optical communication, integrating photonics with traditional electronic chips in a package is becoming more common, enabling faster data transfer and improved signal integrity.

Security Features

- **Hardware-Level Security**—As cyber threats increase, embedding security features directly into semiconductor packages is becoming a trend, especially for applications in defense and secure communications.

These trends will continue to shape the semiconductor packaging landscape, driving innovation and enabling the next generation of electronic devices.

3D Chips Stacking Thermal Management

Stacking chips in integrated circuits, commonly known as 3D IC packaging, presents several thermal challenges. These challenges arise from the increased density of components and the reduced ability for heat dissipation compared to traditional 2D packaging. Below is an overview of the key thermal challenges and strategies to address them:

Thermal Management Challenges

Heat Dissipation

- **Increased Heat Generation—**With multiple layers of active circuitry in a 3D IC, each layer generates heat. The total heat output can be significantly higher than in 2D designs.
- **Reduced Heat Dissipation Pathways—**In 3D stacking, the thermal pathways for heat dissipation are constrained. Heat must travel through multiple layers and interconnects, which can impede effective cooling.

Thermal Gradients

- **Uneven Temperature Distribution—**Different layers of a stacked IC may experience varying temperatures. For instance, the top layers might be hotter than the bottom layers, leading to thermal gradients that can affect performance and reliability.
- **Thermal Mismatch—**Different materials used in the stack (e.g., silicon, solder, and dielectric materials) have different thermal expansion coefficients. This mismatch can cause mechanical stress and potential failure due to thermal cycling.

Reliability Issues

- **Thermal Stress—**The thermal expansion and contraction cycles during operation can induce mechanical stress, leading to potential issues like warping, delamination, or even chip cracking.
- **Thermal Cycling Fatigue—**Repeated heating and cooling cycles can lead to fatigue in interconnects and solder joints, potentially causing reliability issues over time.

Thermal Management Solutions

Advanced Cooling Techniques

- **Heat Spreader/Heat Sink—**Incorporating heat spreaders or heat sinks into the package can help distribute and dissipate heat more effectively. In some designs, these components are integrated into the package itself.
- **Thermal Via Arrays—**Using thermal vias (vertical conductive paths) can help transfer heat from the die to the heat spreader or other cooling solutions more efficiently.

Thermal Interface Materials (TIMs)

- **High-Performance TIMs—**Applying advanced thermal interface materials between the chip layers and heat spreaders can improve heat transfer efficiency. These materials are designed to reduce thermal resistance and enhance heat dissipation.
- **Tim Layer Optimization—**Careful design and optimization of TIM layers are crucial to ensure that they effectively manage the heat between stacked chips.

Die Design and Layout

- **Heat-Aware Design—**Designing the die layout with heat dissipation in mind can help manage thermal challenges. For instance, placing heat-sensitive components away from the hottest regions can mitigate thermal stress.
- **Thermal Redistribution Layers—**Implementing thermal redistribution layers within the stack can help spread heat more uniformly and prevent hot spots.

Material Choices

- **Thermally Conductive Materials—**Using materials with high thermal conductivity for various parts of the stack, such as the die, interconnects, and package substrates, can enhance heat dissipation.
- **Advanced Substrates—**Using advanced substrates with better thermal properties, like silicon-based or specialized high-thermal-conductivity materials, can improve thermal management.

Simulation and Modeling

- **Thermal Analysis Tools—**Employing thermal simulation tools during the design phase can help predict and address potential thermal issues. These tools can model heat flow, temperature distribution, and thermal stresses to optimize the design.
- **Experimental Validation—**Testing prototypes and measuring thermal performance under real operating conditions can provide insights into the effectiveness of the thermal management solutions.

Summary

Stacking chips in ICs introduces significant thermal challenges due to increased heat generation, constrained heat dissipation pathways, and potential reliability issues from thermal stress. Addressing these challenges involves a combination of advanced cooling techniques, effective use of thermal interface materials, careful die design, and material choices. Employing thermal simulation tools and experimental validation helps in optimizing thermal management strategies and ensuring the reliability and performance of stacked ICs.

Micropillars

Micropillars in packaging refer to tiny, pillar-like structures used in advanced packaging technologies for electronics. These micropillars serve several purposes, including enhancing thermal management, improving electrical connectivity, and enabling finer interconnects. Here's a detailed look at micropillars in packaging:

Key Characteristics of Micropillars

1. **Size and Structure**—Micropillars are very small, typically ranging from a few micrometers to hundreds of micrometers in diameter. They are often arranged in a grid or pattern on a substrate or package surface.
2. **Material**—They can be made from various materials, including metals (such as copper or nickel) and polymers. The choice of material depends on the specific requirements of the application, such as thermal conductivity or electrical properties.
3. **Fabrication**—Micropillars are created using advanced fabrication techniques like photolithography, electroplating, or micro-machining. These methods allow for precise control over the size, shape, and spacing of the pillars.

Applications of Micropillars in Packaging

1. **Thermal Management**—Micropillars can enhance heat dissipation by increasing the surface area for heat transfer and improving thermal conductivity. This is crucial in high-performance electronics where managing heat is a key concern.
2. **Electrical Interconnects**—In some packaging designs, micropillars are used as part of the interconnect system to connect different layers or components. They can serve as an alternative to traditional solder bumps or vias, especially in high-density applications.
3. **Mechanical Support**—Micropillars can provide mechanical support and stability for delicate components. They help in reducing stress and strain on the connections and can improve the mechanical robustness of the package.
4. **3D Integration**—Micropillars are used in 3D packaging technologies, where they act as vertical interconnects between different layers or stacked dies. This allows for more compact and efficient designs, enabling higher functionality in smaller footprints.

Advantages of Micropillars

1. **Enhanced Performance**—By improving thermal and electrical performance, micropillars contribute to better overall functionality and reliability of electronic devices.
2. **Miniaturization**—Micropillars support the trend towards smaller, more densely packed electronic components, making them suitable for advanced and compact device designs.
3. **Precision**—The ability to fabricate micropillars with high precision enables the development of high-density, high-performance packaging solutions.

Challenges and Considerations

1. **Manufacturing Complexity**—Creating and integrating micropillars into packaging involves complex manufacturing processes and requires advanced technology, which can increase production costs.
2. **Design Integration**—Effective integration of micropillars into existing packaging designs requires careful consideration of their impact on the overall design, including thermal, electrical, and mechanical factors.
3. **Material Compatibility**—Ensuring compatibility of micropillar materials with other components and processes is crucial to avoid issues related to reliability and performance.

Micropillars represent an advanced approach in packaging technology, addressing the increasing demands for better performance and miniaturization in electronics. Their use continues to evolve with advancements in manufacturing techniques and applications.

Fabrication of micropillars is detailed in Appendix H.

Additional Considerations

- **Material Selection—**Choosing the right materials for micropillar fabrication is crucial for ensuring desired performance characteristics, such as thermal conductivity, electrical conductivity, and mechanical strength.
- **Precision and Tolerance—**Achieving high precision and tight tolerances in micropillar fabrication is essential for high-performance applications and ensuring reliable operation.

These steps outline the general process for micropillar fabrication in packaging. The specific details and techniques may vary depending on the application, materials used, and the desired characteristics of the final product.

Trends in Micropillars' Spacing

The trend in pillar spacing within semiconductor packaging is closely tied to the broader industry drive towards miniaturization and enhanced performance. Here are the key trends:

Reduction in Pillar Spacing

- **Decreasing Pitch**—Pillar spacing, like bump pitch, has been steadily decreasing over time. This reduction allows for higher interconnect density, which is crucial for accommodating more connections in increasingly smaller chip sizes.
- **Current State (2024)**—Typical micropillar pitches in advanced packaging technologies, such as flip-chip and 2.5D/3D ICs, are in the range of 40-100 µm.
- **Trends**—The industry is moving towards finer pitches, with expectations of pitches as low as 10-20 µm by 2030. This reduction in pitch is essential to support higher interconnect densities, which are crucial for advanced node integration, high-performance computing, and AI applications.
- **Finer Features**—As devices continue to shrink and performance requirements increase, the industry has moved from larger pitch sizes (e.g., 250-300 µm) to much finer pitches (e.g., 60-80 µm).

Advanced Manufacturing Techniques

- **Precision Engineering**—New manufacturing techniques, such as photolithography and advanced etching methods, have enabled the production of pillars with finer spacing.
- **Improved Materials**—The development of new materials for pillars and interconnects has also supported the reduction in spacing, ensuring reliability and performance even at smaller scales.

Impact of 3D Integration

- **3D Packaging**—The shift towards 3D integration, where chips are stacked vertically, has also influenced pillar spacing. Tighter pillar spacing is necessary to support the increased density of interconnections in these advanced packaging formats.

Challenges and Solutions

- **Signal Integrity**—As pillar spacing decreases, maintaining signal integrity becomes more challenging. Innovations in design and materials are addressing these issues.
- **Thermal Management**—Closer pillar spacing can lead to heat management challenges, which are being addressed through advanced cooling solutions and thermal interface materials.

Future Directions

- **Continued Miniaturization**—The trend towards smaller pillar spacing is expected to continue as the demand for high-performance, compact devices grows.
- **Integration with Emerging Technologies**—As technologies like AI, IoT, and 5G continue to evolve, the need for finer pillar spacing in packaging will remain critical to meet the performance and size requirements of next-generation devices.

Overall, the trend in pillar spacing in semiconductor packaging is towards increasingly finer pitches, driven by the need for higher performance and smaller, more efficient devices. See Figure 11.

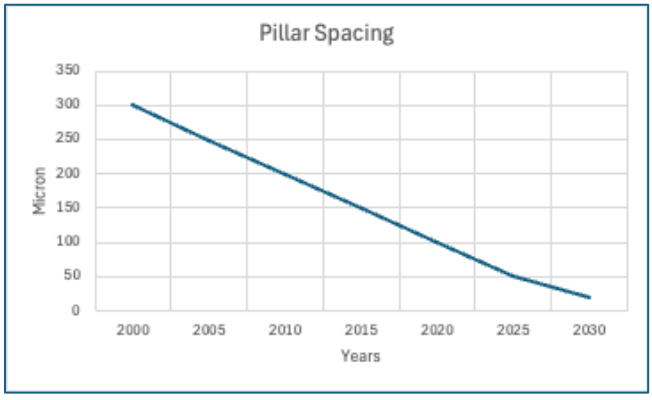


Figure 11. Historical trend of pillar spacing and future forecast

Flip Chip and Micropillars

The differences between flip-chip and wafer-level pillars in semiconductor packaging are primarily related to their design, application, and manufacturing processes. Here's an overview:

Flip-Chip Pillars

- **Design and Structure**—In flip-chip packaging, the pillars (also referred to as bumps) are metallic structures, typically made of solder or copper, that connect the die (chip) to the substrate. These pillars are located on the active surface of the die, which is then "flipped" to attach to the substrate.
- **Application**—Flip-chip technology is widely used in high-performance applications, such as processors, GPUs, and other complex integrated circuits where high interconnect density and performance are required.
- **Manufacturing Process**:
  - **Deposition**—The pillars are created on the chip's surface through processes like electroplating or solder deposition.
  - **Reflow Soldering**—The chip is flipped and placed onto the substrate, where the pillars are connected through a reflow soldering process, creating strong and reliable electrical connections.
- **Advantages**:
  - High interconnect density.
  - Improved electrical and thermal performance due to shorter interconnect paths.

    - Greater miniaturization potential.
- **Challenges**:
    - More complex and costly manufacturing process.
    - Requires precise alignment during assembly.2. Wafer-Level Pillars

Wafer-level Pillars

- **Design and Structure**—In wafer-level packaging (WLP), the pillars are also metallic structures but are created at the wafer level before the individual dies are separated. This process allows the entire wafer to be packaged simultaneously.
- **Application**—WLP is often used for smaller, lower-power devices like sensors, RF components, and some mobile and wearable devices. It's favored where small size and cost-effectiveness are critical.
- **Manufacturing Process**:
    - **Batch Processing**—The pillars are formed directly on the wafer in a batch process, which includes steps like photolithography, electroplating, and etching.
    - **Singulation**—After the pillars are formed, the wafer is diced into individual chips that are already packaged.
- **Advantages**:
    - Cost-effective for high-volume production.
    - Simplified manufacturing and testing process since packaging is done at the wafer level.
    - Ideal for small, lightweight, and low-power applications.
- **Challenges**:
    - Limited to applications with lower power and performance requirements compared to flip-chip.
    - Less flexibility in design and customization.

Summary of Key Differences

- **Application**—Flip-chip is used for high-performance, high-density applications, while wafer-level packaging is used for smaller, cost-sensitive devices.
- **Manufacturing**—Flip-chip involves packaging each die individually, while WLP processes the entire wafer at once.
- **Cost and Complexity**—Flip-chip packaging is generally more complex and costly but offers superior performance. WLP is more cost-effective and suitable for mass production of smaller devices.

Each approach has its strengths, depending on the specific needs of the device and application. Figure 12 shows historical trends and the future forecast of bump and pillar spacing.

Through Silicon Vias

Through-silicon vias are a critical technology in advanced semiconductor packaging, particularly for enhancing performance and enabling the integration of complex systems. Below is an overview of the function, fabrication, and packaging aspects of TSVs:

Vertical Interconnects—TSVs are vertical electrical connections that pass through a silicon wafer or die. They enable electrical communication between different layers of silicon or between different chips in a stacked configuration.

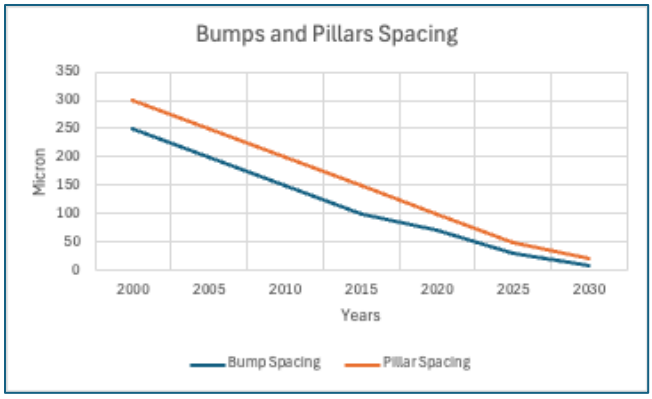


Figure 12. Historical trend of Bump and Pillar spacing and future forecast

This vertical interconnectivity helps to significantly reduce the distance signals need to travel, leading to faster data transfer rates and lower power consumption.

Enhanced Bandwidth and Performance—By providing high-bandwidth communication paths between different layers or chips, TSVs improve the overall performance of the system. They reduce latency and increase data throughput compared to traditional lateral interconnects, such as those found in conventional 2D packaging.

Integration and Miniaturization—TSVs facilitate the integration of multiple functionalities into a smaller footprint. They enable 3D stacking of chips, which allows for more compact and efficient designs. This is particularly useful in applications where space is at a premium, such as in mobile devices and high-performance computing.

The fabrication process for TSVs is detailed in Appendix I.

Packaging with TSVs

- **Die Stacking**—In 3D packaging, TSVs allow for the stacking of multiple dies on top of each other. This 3D integration can be used to combine different functionalities, such as memory and logic, into a single package. TSVs connect the different layers of stacked dies, enabling high-speed communication between them.
- **Interposer Technology**—TSVs are also used in interposer technology, where a silicon interposer with TSVs is placed between the main die and the package substrate. The interposer acts as a high-bandwidth bridge, facilitating communication between the die and other components, such as memory or high-speed I/O interfaces.
- **Thermal Management**—TSVs can help with thermal management by distributing heat more evenly across the package. However, managing heat in 3D stacks remains a challenge, and additional thermal dissipation techniques may be required.
- **Assembly and Testing**—The final packaging step involves bonding the TSV-equipped die or interposer to the package substrate. This is typically done using techniques such as solder bump bonding or wafer-level bonding. After assembly, the packaged device undergoes testing to ensure functionality and performance.

The Benefits and Challenges for TSV Technology

Overall, TSVs represent a significant advancement in semiconductor packaging technology, enabling higher performance, better integration, and more compact designs. See Table 18.

Table 18. TSV Benefits and Challenges

Benefit	Description
Increased Bandwidth	• TSVs provide high-speed data transfer capabilities due to shorter signal paths.
Reduced Power Consumption	• By minimizing the distance signals travel, TSVs reduce power dissipation.
Space Efficiency	• 3D stacking and integration of multiple components in a single package save board space.

Challenge	Description
Complexity	• The TSV fabrication process is complex and requires precise control over multiple steps.
Cost	• TSV technology can be expensive due to the advanced equipment and materials required.
Thermal Management	• Efficiently managing heat in 3D stacked structures is challenging and requires innovative solutions.

TSV Application in 3D stacking

Through-Silicon Via (TSV) technology is a key component in 3D IC (integrated circuit) packaging, enabling vertical stacking of dies and interconnecting them through vias that pass through the silicon wafer. The spacing between TSVs is a critical factor that impacts performance, cost, and reliability.

Trends in TSV Attributes

Decreasing TSV Spacing

- **Higher Interconnect Density**—As the demand for higher interconnect density in 3D ICs grows, TSV spacing has been steadily decreasing. This allows more TSVs to be integrated into a given area, enhancing the performance and functionality of the IC.
- **Smaller Form Factors**—The trend towards miniaturization in electronics pushes for reduced TSV spacing, contributing to smaller and more compact devices.

Advances in Fabrication

- **Precision Etching**—Improvements in etching techniques have enabled the creation of TSVs with smaller diameters and tighter spacing.
- **Material Innovations**—New materials and deposition methods have allowed for more precise control over TSV dimensions, further reducing the spacing between them.

Challenges with Reduced Spacing

- **Thermal Management**—As TSVs are placed closer together, managing the heat generated within the IC becomes more challenging. This has led to innovations in thermal management techniques.
- **Signal Integrity**—Reduced TSV spacing can lead to increased crosstalk and signal interference, necessitating advanced design and materials to mitigate these issues.

Future Directions

- **Integration with Advanced Technologies**—TSVs are increasingly being used in combination with other advanced packaging techniques like 2.5D and 3D integration, further pushing the limits of TSV spacing.
- **Continued Miniaturization**—The trend towards finer TSV spacing is expected to continue as the industry moves towards even smaller and more powerful devices.

Historical Trends in TSV Spacing

Through-silicon vias are critical components in 3D integrated circuits and advanced packaging technologies. They enable vertical electrical connections through silicon wafers or dies, allowing for the stacking of multiple layers of semiconductor devices. The spacing of TSVs has evolved significantly over the years, driven by the need for higher performance, increased integration density, and the development of new applications. Here's a look at the historical trends in TSV spacing leading up to 2030:

Early Development (2000s)

- **TSV Spacing**—20 µm to 100 µm
- **Context**—In the early 2000s, TSV technology was in its infancy. The focus was primarily on proof-of-concept and early-stage development, with relatively large TSVs and wider spacing. The main applications were memory stacking and the initial exploration of 3D ICs.

Initial Commercialization (2010s)

- **TSV Spacing**—10 µm to 50 µm
- **Context**—During the 2010s, TSVs became more commercially viable, especially in high-performance applications like memory (e.g., High Bandwidth Memory or HBM) and some early 3D ICs. The spacing between TSVs began to decrease as the industry improved its understanding of the technology and manufacturing capabilities advanced.
- **Applications**—DRAM stacking (e.g., HBM), 2.5D packaging with interposers, image sensors.

Advanced Packaging and 3D Integration (2020s)

- **TSV Spacing**—5 µm to 10 µm
- **Context**—In the 2020s, TSVs became increasingly critical in high-performance computing, AI accelerators, and other advanced applications. The push for higher density and performance led to further reductions in TSV spacing, with pitches as small as 5 µm becoming common in leading-edge designs.
- **Applications**—3D NAND, HBM2/2E/3, advanced processors with 3D stacking, 2.5D ICs with fine-pitch interposers.

Emerging Trends and Projections to 2030

- **TSV Spacing (2020s to 2030)**—1 µm to 5 µm
- **Context**—As we approach 2030, TSV technology is expected to continue its evolution, with pitches shrinking further to 1-5 µm. This trend is driven by the increasing demand for even higher integration densities, lower power consumption, and the need for more efficient thermal management in 3D ICs.
- **Applications**—By 2030, TSVs are expected to be widely used in advanced 3D chiplet architectures, heterogeneous integration, and high-performance computing. Applications like AI/ML processors, data centers, and advanced memory technologies will drive the adoption of these fine-pitch TSVs.
- **Innovations**—The introduction of new materials (e.g., low-k dielectrics for insulation), hybrid bonding technologies, and the development of more precise manufacturing techniques will facilitate this trend. Additionally, improvements in lithography (e.g., EUV) and etching techniques will allow for tighter TSV pitches.

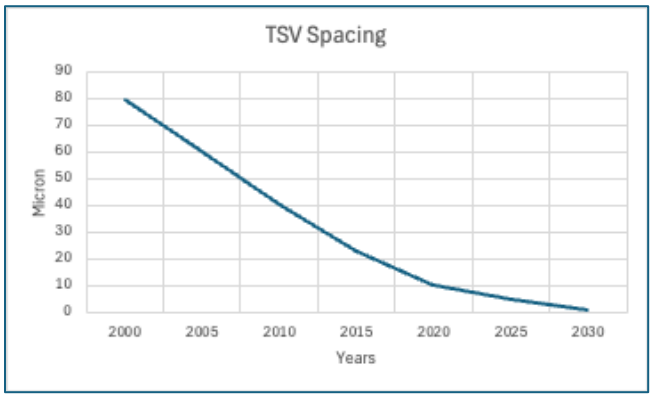


Figure 13. Historical trends in TSV spacing and future forecast

Key Factors Influencing TSV Spacing Trends

- **Performance Demands**—The need for higher bandwidth, lower latency, and more efficient power delivery in 3D ICs is a major driver for reducing TSV spacing.
- **Thermal Management**—As TSVs become more densely packed, managing the heat generated in 3D stacks becomes more challenging. Innovations in thermal management techniques are essential to supporting this trend.
- **Manufacturing Technology**—Advances in lithography, etching, and deposition technologies will enable the continued reduction in TSV spacing.
- **Cost Considerations**—While finer TSV pitches offer performance benefits, they also increase manufacturing complexity and cost. By 2030, the industry will need to balance these factors to achieve widespread adoption of fine-pitch TSVs.

Conclusion

- 2000s—TSV spacing started large, around 20-100 µm, with initial research and development efforts.
- 2010s—Spacing decreased to 10-50 µm as TSVs entered commercial use in memory stacking and early 3D ICs.
- 2020s—Further reduction to 5-10 µm, driven by advanced packaging and high-performance applications.
- 2030 Projection—Expect TSV spacing to shrink to 1-5 µm, with widespread use in 3D chiplets, high-performance computing, and other cutting-edge applications.

Micropillars and TSV roles in SiP

System-in-Package technology integrates multiple components, such as processors, memory, sensors, and passive elements, into a single package. This approach enables greater functionality in a compact form factor, making it crucial for modern electronic devices. Two key technologies—micropillars and through-silicon vias—play vital roles in the development and functionality of SiP. Here's an overview of their roles:

Micropillars in SiP

Interconnect Density and Precision

- **High-Density Interconnects**—Micropillars are small metallic structures used to create dense interconnections between different dies or components within a SiP. They enable tight spacing and high interconnect density, which is essential for achieving the compact size and high performance of SiP.
- **Precision Alignment**—Micropillars allow for precise alignment of dies during the assembly process. This precision ensures that the connections between components are reliable, and that signal integrity is maintained, which is critical for the performance of the SiP.

Electrical and Thermal Conductivity

- **Efficient Signal Transmission**—Micropillars facilitate efficient electrical signal transmission between components within the SiP. By minimizing the distance between interconnected elements, they reduce signal loss and latency, enhancing the overall performance of the system.
- **Thermal Management**—In addition to their role in electrical connectivity, micropillars can also help with thermal conductivity. They provide a path for heat dissipation from one die to another or to the package substrate, aiding in the thermal management of the SiP.

Manufacturing and Assembly

- **Fine-Pitch Interconnections**—The small size of micropillars allows for fine-pitch interconnections, which are essential in SiP where space is at a premium. This enables more components to be packed into a smaller area, which is particularly important in mobile and wearable devices.
- **Compatibility with Flip-Chip Bonding**—Micropillars are often used in flip-chip bonding, a process where the die is flipped and attached directly to the package substrate using the pillars. This method improves electrical performance and reduces the overall footprint of the SiP.

TSVs in SiP

Vertical Interconnection

- **3D Stacking**—TSVs are vertical electrical connections that pass through the silicon wafer, allowing for 3D stacking of dies. In SiP, TSVs enable the integration of multiple layers of active and passive components, significantly increasing the functionality and performance of the package.
- **High Bandwidth and Low Latency**—By providing direct vertical connections between stacked dies, TSVs reduce the length of interconnects, leading to higher data bandwidth and lower latency. This is particularly important for SiP applications that require fast data processing and communication, such as in AI and high-performance computing.

Power Delivery and Thermal Management

- **Efficient Power Distribution**—TSVs play a crucial role in distributing power across different layers of the SiP. By providing low-resistance paths, they ensure efficient power delivery to all components, which is vital for maintaining the performance and reliability of the system.
- **Heat Dissipation**—Like micropillars, TSVs can also help with heat dissipation in a 3D SiP. They provide additional pathways for heat to escape from the inner layers of the stacked dies to the outer surfaces, helping to manage thermal challenges in densely packed systems.

Integration of Diverse Technologies

- **Heterogeneous Integration**—TSVs facilitate the integration of diverse technologies within a single SiP. For example, different types of memory, processors, and specialized sensors can be stacked and interconnected using TSVs, allowing for a multifunctional package that meets specific application needs.
- **Advanced Packaging Techniques**—TSVs are often used in conjunction with advanced packaging techniques like 2.5D and 3D integration. This enables the combination of multiple chips with different process technologies, optimizing performance, power, and cost.

Synergy Between Micropillars and TSVs in SiP

- **Complementary Roles**—Micropillars and TSVs complement each other in SiP technology. ***While micropillars provide high-density horizontal interconnections at the die level, TSVs offer vertical connectivity, enabling complex 3D architectures***.
- **Enhanced Performance and Miniaturization**—Together, these technologies allow for greater miniaturization without sacrificing performance. This is particularly important in applications like smartphones, IoT devices, and wearable electronics, where space is limited, but high performance is required.
- **Scalability**—The use of micropillars and TSVs makes it possible to scale SiP designs to meet the demands of various applications, from consumer electronics to industrial and automotive systems.

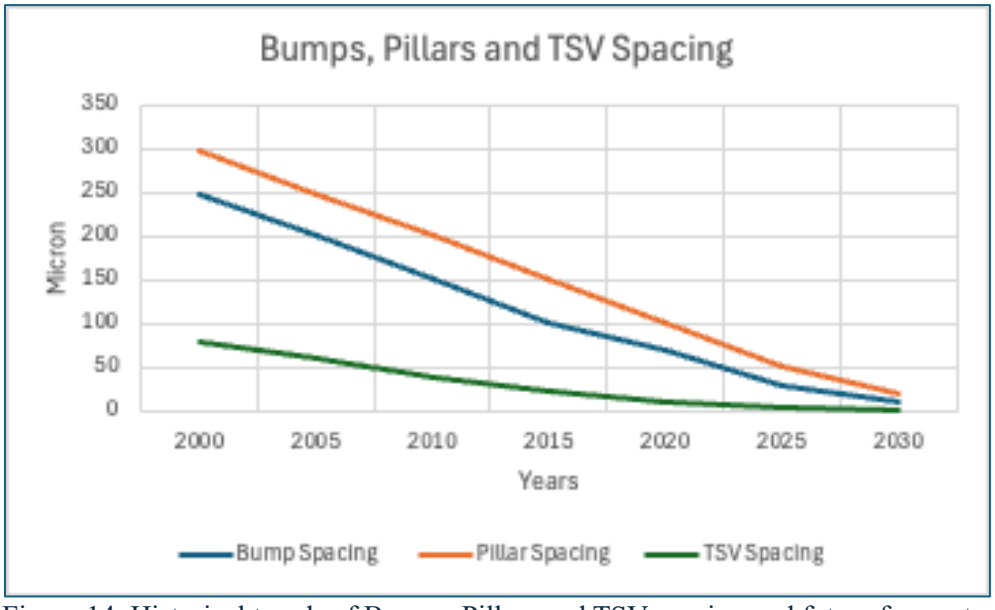
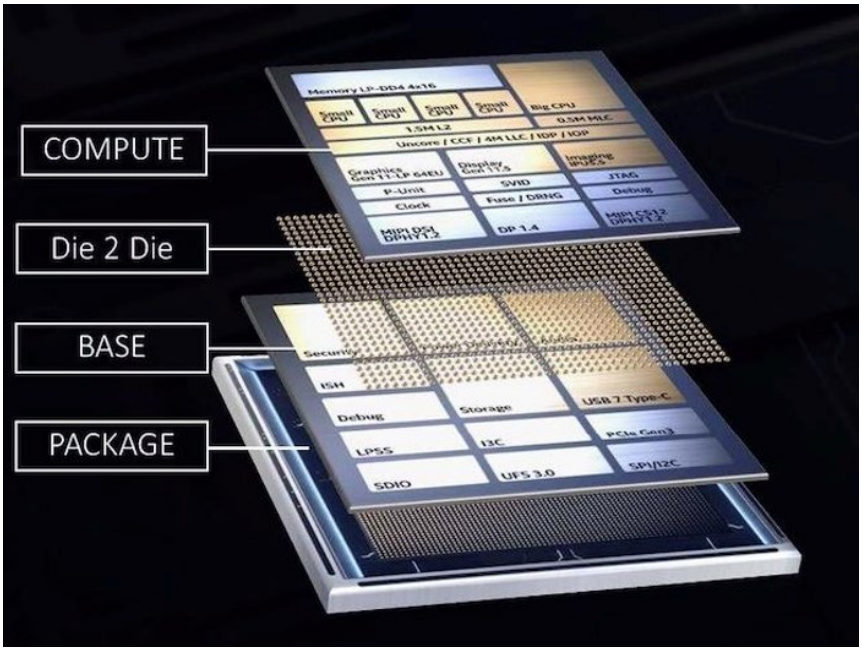


Figure 14. Historical trends of Bumps, Pillars and TSV spacing and future forecast

Conclusion

Micropillars and TSVs are integral to the success of System-in-Package technology. Micropillars enable high-density, fine-pitch interconnections with efficient thermal management, while TSVs provide the vertical interconnects necessary for 3D stacking and heterogeneous integration. Together, they support the miniaturization, performance, and functionality required in modern electronic systems, making SiP a powerful solution for the next generation of semiconductor devices.



Source: Intel

Figure 15. Schematic example of 3D integration in SiP

Sixth ETP Summary—System in Package

The development and adoption of silicon-based System-in-Package technologies are driven by various factors that focus on enhancing the integration of multiple components, improving performance, reducing size, and enabling advanced functionalities. Silicon SiP technology is increasingly important in applications where traditional single-chip solutions are insufficient, particularly in advanced computing, telecommunications, consumer electronics, and automotive industries. Semiconductor packaging is evolving rapidly due to the increasing demands for smaller, faster, and more energy-efficient electronic devices. The most common SiP implementation is based on flip chip technology. In addition, micropillars and TSV technology modules have been developed to further increase packing density, enhance connectivity and reduce signal propagation delays; trends show continuous reduction in bumps, micropillars and TSV pitch.

7. SiP Role in Reducing Signal Propagation Delay (SPD)

System-in-Package technology plays a significant role in reducing **signal propagation delay**, which is crucial for enhancing the performance and efficiency of modern electronic systems. Signal propagation delay refers to the time it takes for a signal to travel from one point to another within a circuit or system. Reducing this delay is essential for high-speed computing, communication, and processing tasks.

Table 19. SiP Technology Benefits in SPD Reduction

Attribute	Benefits
Shorter Interconnects	• Reduced Physical Distance: In SiP, multiple components such as processors, memory, and sensors are integrated into a single package. This proximity reduces the physical distance that signals need to travel between components, which directly reduces signal propagation delay. • Minimized Parasitic Effects: Shorter interconnects within the SiP also mean reduced parasitic capacitance, inductance, and resistance. These parasitics can slow down signal transmission, so minimizing them helps improve signal speed.
High-Density Interconnects	• Dense Integration: SiP allows for high-density integration of components, using technologies like micropillars and fine-pitch interconnects. The increased density enables more direct and shorter pathways for signals, further reducing the time it takes for signals to propagate. • Advanced Packaging Techniques: Techniques like flip-chip bonding used in SiP ensure that connections between the die and substrate are as short and direct as possible, reducing delay.
3D Integration and TSVs	• Vertical Stacking: SiP can incorporate through-silicon vias for vertical interconnection in 3D stacked architectures. TSVs enable signals to travel vertically between layers, reducing the need for long, complex horizontal routing. This vertical integration dramatically reduces signal propagation delay compared to traditional 2D layouts. • Reduced Crosstalk: By enabling direct vertical connections, TSVs also reduce crosstalk between signals, which can otherwise cause delays and signal degradation in densely packed systems.
Improved Signal Integrity	• Lower Signal Loss: SiP technology improves signal integrity by reducing the length of interconnects, which lowers the chances of signal loss or degradation. Better signal integrity means that signals can travel faster and with less delay. • Optimized Materials and Design: The materials and design techniques used in SiP are optimized to support high-frequency signals with minimal delay. For example, low-k dielectrics and advanced substrate materials can be used to reduce capacitance and improve signal speed.
Integration of Heterogeneous Components	• Close Proximity of Components: SiP allows for the integration of heterogeneous components, such as processors, memory, and RF modules, within the same package. This proximity ensures that signals between different types of components experience minimal delay, which is especially important in high-performance computing and communication systems. • Optimized Signal Pathways: SiP design can be optimized to create the most efficient signal pathways between integrated components,

Attribute	Benefits
	reducing the time it takes for signals to travel from one component to another.
Reduced Power Consumption	• Lower Power, Faster Signals: SiP can help reduce overall power consumption by shortening interconnects and reducing the need for signal boosting or amplification. Lower power consumption contributes to faster signal transmission, as signals encounter less resistance and loss. • Efficient Power Distribution: SiP also allows for more efficient power distribution, ensuring that signals receive consistent power levels throughout the package, reducing variations that could lead to delays.
Enhanced System Performance	• Faster Data Processing: By reducing signal propagation delay, SiP enhances the overall speed and performance of electronic systems. This is particularly important in applications such as high-performance computing, where rapid data processing and communication are critical. • Real-Time Communication: In communication systems, reduced signal delay means faster data transmission and reception, which is vital for real-time applications like autonomous vehicles, 5G networks, and IoT devices.

Conclusion

SiP technology is instrumental in reducing signal propagation delay by enabling shorter, more direct interconnects, improving signal integrity, and allowing for the integration of heterogeneous components in close proximity. These benefits lead to faster data processing, enhanced system performance, and more efficient power consumption, making SiP a key enabler of high-speed, high-performance electronic systems.

Historical Trends of Signal Propagation Delay and Interconnect Density in SiP

The historical trends in signal propagation delay and interconnect density in System-in-Package technologies reflect the broader evolution of semiconductor packaging, driven by the demand for higher performance, greater functionality, and more compact designs. SiP integrates multiple semiconductor dies and passive components into a single package, allowing for increased system complexity while maintaining a small form factor.

Signal Propagation Delay Historical Trends

Early Stages (1990s - 2000s)

- **Signal Propagation Delay**—Relatively high, typically in the range of **hundreds of picoseconds per millimeter.**
- **Context**—Early SiP designs were often based on wire bonding techniques, leading to relatively long interconnect paths between components. This contributed to higher signal propagation delays, as the signal had to traverse longer distances and experience more parasitic effects, such as inductance and capacitance.
- **Key Technologies**—Wire bonding, traditional substrate materials like FR-4.

Advancements and Transition (2010s)

- **Signal Propagation Delay**—Reduction to **tens of picoseconds per millimeter**.
- **Context**—The adoption of flip-chip technology and the use of more advanced substrate materials (like BT resin and ABF) allowed for shorter interconnects and better control over parasitic effects. Additionally, the integration of passive components within the package itself helped reduce overall signal delay.
- **Key Technologies**—Flip-chip, advanced substrates, use of high-speed signal routing techniques.

Recent Developments (2020s)

- **Signal Propagation Delay**—Continued reduction to **low tens of picoseconds per millimeter**.
- **Context**—With the advent of advanced packaging techniques like embedded wafer-level ball grid array (eWLB) and fan-out wafer-level packaging (FOWLP), SiP designs have further minimized signal propagation delay. These technologies allow for much finer interconnects and closer proximity between components, significantly reducing signal delay.
- **Key Technologies**—FOWLP, eWLB, and integration of high-speed interconnects like copper pillars and microbumps.

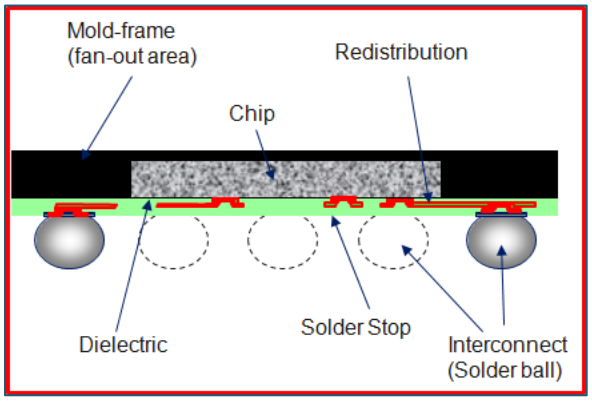


Figure 16. Key components contributing to SPD in modern packages

Projections to 2030

- **Signal Propagation Delay**—Potential reduction to single-digit picoseconds per millimeter.
- **Context**—As SiP technology continues to evolve, further reductions in signal propagation delay are expected. This will be driven by the adoption of even more advanced materials, such as low-k dielectrics and high-conductivity metals, as well as innovations in 3D packaging and interconnect techniques like hybrid bonding.
- **Key Technologies**—3D packaging, hybrid bonding, further advancements in low-k and ultra-low-k dielectrics.

Interconnect Density Historical Trends

Early Stages (1990s - 2000s)

- **Interconnect Density**—Low, with typical interconnect pitches in the range of 100 µm or more.
- **Context**—Early SiP technologies were primarily focused on integrating a few large components, with relatively coarse interconnect pitches. The interconnect density was limited by the capabilities of the available packaging technologies, such as wire bonding and conventional PCB substrates.
- **Key Technologies**—Wire bonding, FR-4 substrates.

Advancements and Transition (2010s)

- **Interconnect Density**—Moderate, with interconnect pitches reducing to 40-100 µm.
- **Context**—The transition to flip-chip and other advanced packaging techniques allowed for a significant increase in interconnect density. The use of finer substrates and improved lithography techniques enabled tighter spacing between interconnects, allowing for more complex SiP designs.
- **Key Technologies**—Flip-chip, finer pitch interposers, advanced substrates like BT resin and ABF.

Recent Developments (2020s)

- **Interconnect Density**—High, with pitches shrinking to 10-40 µm.
- **Context**—The adoption of fan-out and fan-in wafer-level packaging, along with the move towards 2.5D and 3D IC integration within SiP, has led to a substantial increase in interconnect density. These techniques

enable very fine interconnects, supporting the integration of multiple dies and passive components in a highly compact form factor.

- **Key Technologies**—FOWLP, 2.5D/3D ICs, ultra-fine pitch interposers, copper pillar, and microbump technologies.

Projections to 2030

- **Interconnect Density**—Very high, with pitches potentially reaching 5-10 µm or less.
- **Context**—By 2030, interconnect density in SiP is expected to reach new heights, driven by advances in 3D stacking, chiplet integration, and further miniaturization of interconnects. Technologies like hybrid bonding and the use of advanced materials will allow for extremely dense integration, enabling more functionality in even smaller packages.
- **Key Technologies**—Hybrid bonding, 3D ICs, sub-10 µm pitch interconnects, advanced substrates with embedded passive components.

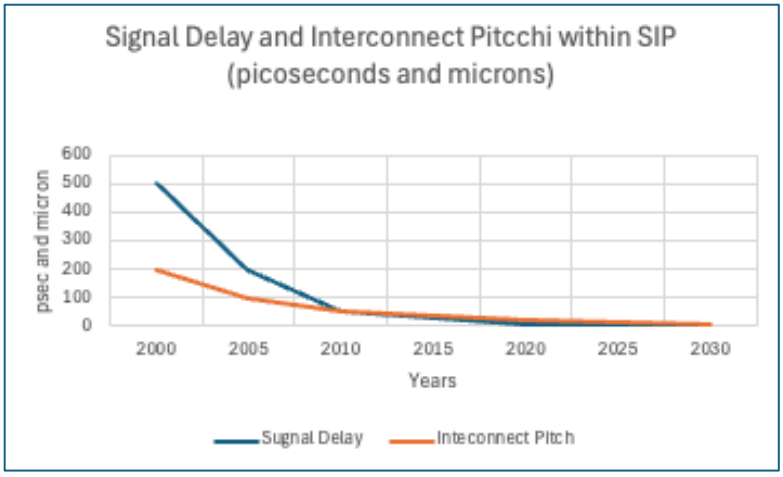


Figure 17. Relation between SPD and interconnect density

Summary of SPD and Interconnect Density Trends

- **Signal Propagation Delay**—
  - From hundreds of picoseconds per millimeter in the 1990s-2000s to single-digit picoseconds per millimeter projected by 2030.
  - Driven by shorter interconnects, advanced materials, and 3D integration.
- **Interconnect Density**:
  - From pitches over 100 µm in the 1990s-2000s to potentially sub-10 µm by 2030.
  - Driven by innovations in advanced packaging, 3D stacking, and finer lithography techniques.

These trends reflect the ongoing push for greater integration, performance, and miniaturization in semiconductor packaging, as SiP technologies continue to evolve and play a critical role in next-generation electronics.

Motivation and Roles of Capacitors and Inductors in System-in-Package and Board Assembly

Capacitors and Inductors are fundamental in ensuring the reliable operation of modern electronic systems, enabling high performance and stable operation in complex SiP and PCB designs.

Table 20. Use of Capacitors and Inductors in SiP

	Motivation	Roles
Capacitors	• Power Integrity: Capacitors are critical in maintaining power integrity across the SiP and PCB. They help in stabilizing the power supply by providing a local reservoir of charge, which helps to smooth out fluctuations and supply clean power to various components. • Signal Integrity: Capacitors also play a role in signal conditioning by filtering out noise and reducing electromagnetic interference (EMI), which is essential for maintaining the quality of high-speed signals.	• Decoupling and Bypass: Capacitors are commonly used for decoupling and bypassing in SiP and PCB assemblies. They decouple the power supply from the rest of the circuit by filtering out high-frequency noise and providing a low-impedance path to ground for AC signals. • Filtering: In signal processing, capacitors are used in filtering circuits to allow certain frequencies to pass while blocking others. This is crucial in both analog and digital circuits to ensure that only the desired signals are processed. • Energy Storage: Capacitors temporarily store energy, which can be quickly released when needed. This is particularly important in high-speed digital circuits where sudden changes in current demand can occur.
	Motivation	**Roles**
Inductors	• Energy Storage in Magnetic Fields: Inductors store energy in their magnetic fields and are used in circuits where energy needs to be stored temporarily and released slowly, which is particularly important in power management and voltage regulation. • Signal Conditioning: Inductors help in filtering and conditioning signals, particularly in RF and high-frequency applications, where they block high-frequency noise while allowing the desired signals to pass through.	• Power Regulation: In DC-DC converters and voltage regulators, inductors work alongside capacitors to smooth out the voltage and current supplied to the components in the SiP or PCB. They help in reducing voltage ripple and improving overall power efficiency. • EMI Filtering: Inductors are used in combination with capacitors to create filters that suppress electromagnetic interference. This is essential in SiP and PCB designs, especially in environments with significant RF noise. • Impedance Matching: In high-frequency circuits, inductors are used to match the impedance between different stages of the circuit, which is crucial for maximizing power transfer and minimizing signal reflections.

Conclusion

Capacitors are essential for ensuring power and signal integrity, energy storage, and noise filtering in SiP and board assemblies. Inductors are key to power regulation, EMI suppression, and signal conditioning, particularly in high-

frequency applications. Together, these components are fundamental in ensuring the reliable operation of modern electronic systems, enabling high performance and stable operation in complex SiP and PCB designs.

Some practical examples of actual SiP packaging technologies

SiP (System In Package) is an advanced semiconductor packaging technology that represents a significant evolution in integrated circuit packaging, designed to improve performance, integration density, and thermal management. SiP integrates multiple chiplets or dice into a single package using advanced 3D packaging techniques. This allows combining different types of chips (such as logic, memory, and analog) in one package to create a more compact and efficient system. **SiP** typically uses a silicon interposer or a similar substrate layer to facilitate the high-density interconnections between the co-packaged chips. This approach educes power consumption by minimizing the length of interconnects and improving signal integrity.

There are several packaging technologies, chip architectures and systems that provide solutions for building and implementing SiPs. These include, for example, TSMC's chip-on-wafer-on-substrate (CoWoS) and Integrated Fan-Out (InFO), AMD's EPYC processors' technologies, and Intel's Ponte Vecchio, a high-performance compute GPU designed primarily for the exascale computing market, their EMIB (Embedded Multi-die Interconnect Bridge) and Foveros packaging technologies.

The examples are provided in Appendix J for tutorial purposes only and do not represent recommendations or endorsements from ETP. For more detailed information it is recommended that the companies are directly contacted

Seventh ETP Summary—Signal Propagation Delay

System-in-Package technology plays a significant role in reducing **signal propagation delay**, which is crucial for enhancing the performance and efficiency of modern HPC electronic systems. Leading electronics companies have developed customized versions of SiP, but they all leverage minimization of space between components. The packaging technology is adapting on step at a time to this new active role, but it is still dominated by the traditional approach of minimizing cost of packaging. For instance, interconnections of reduced length running on the organic substrate between adjacent dice within the package provide better signal performance than interconnections running from package to package on the PCB but do not provide the best possible reduction in signal propagation.

To alleviate this problem technologies like EMIB have been developed to provide high-density, high-bandwidth interconnect technology that enables efficient communication between different chiplets on a single package without the need for large interposers or complex wiring schemes. This involves embedding a small silicon bridge directly into the substrate of the processor package. This bridge connects the different chiplets, allowing them to communicate with each other with low latency and high bandwidth. It would seem simpler to just replace the organic package substrate all together with a silicon substate but traditional cost consideration have led to this somewhat "creative" solution. However, going forward a different approach driven by the goal of achieving the ultimate performance may be appropriate.

8. Dynamic Power Consumption in ICs and Its Impact on Temperature and Leakage

Dynamic power consumption in ICs is a critical factor in modern chip design, especially as the number of transistors, operating frequency, and supply voltage increase. This sequence of increasing power consumption leads to higher IC temperatures, which in turn increases leakage currents. Here's a step-by-step outline of this process, including relevant equations and graphs.

Dynamic Power Consumption

- Basic Formula:

$$P_{dynamic} = \alpha \cdot C_L \cdot V_{DD}^2 \cdot f$$

  - α: Activity factor (fraction of transistors switching per clock cycle).
  - C_L: Load capacitance.
  - V_{DD}: Supply voltage.
  - f: Operating frequency.
- Explanation:
  - Number of Transistors: As the number of transistors increases, the total load capacitance C_L also increases, leading to higher dynamic power consumption.
  - Frequency: Increasing the operating frequency f linearly increases dynamic power consumption.
  - Voltage: The dynamic power consumption has a quadratic dependence on the supply voltage V_{DD}, meaning even a small increase in voltage significantly increases power consumption.

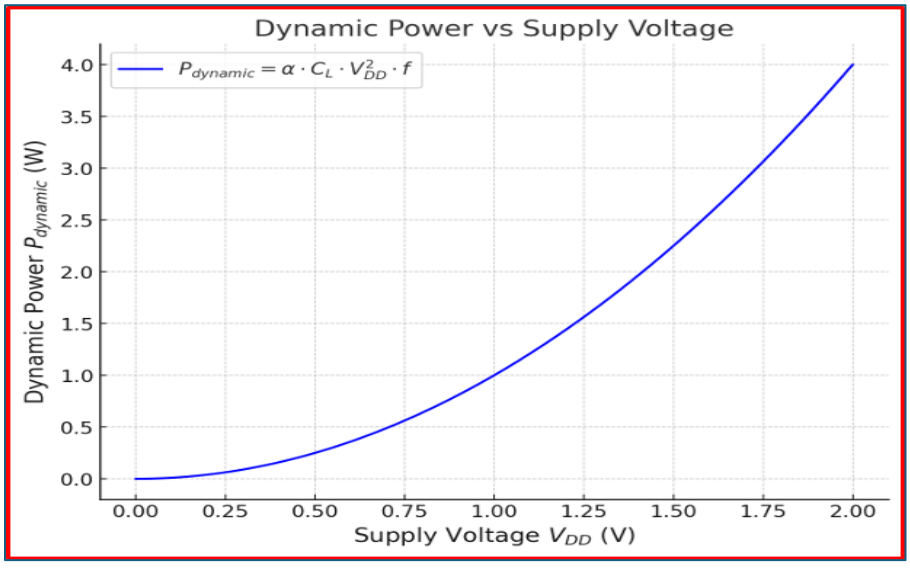


Figure 18. Dynamic power consumption as a function of supply voltage

Temperature Increase:

- Thermal Equation:

$$\Delta T = \theta_{JA} \cdot P_{total}$$

  - P_{total}= Total power dissipation in the IC results in heat generation. As dynamic power $P_{dynamic}$ increases, the IC's temperature ΔT rises proportionally.

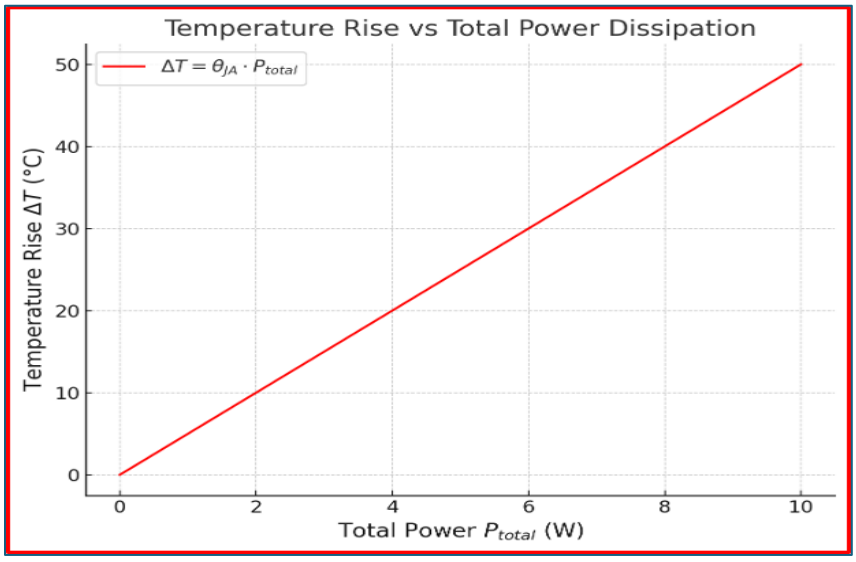


Figure 19. Temperature increase as a function of power

3. Increase in Leakage Power:

- $I_{leakage} = I_0 \cdot e^{-\frac{q \cdot V_{th}}{k \cdot T}}$

  - Io: Constant depending on the technology.
  - Vth: Threshold voltage.
  - T: Temperature.
  - q: Charge of an electron.
  - k: Boltzmann constant.

Explanation:

- As the temperature T increases, the leakage current $I_{leakage}$ also increases exponentially.
- This increase in leakage current leads to higher leakage power consumption: $P_{leakage}=VDD \cdot I_{leakage}$

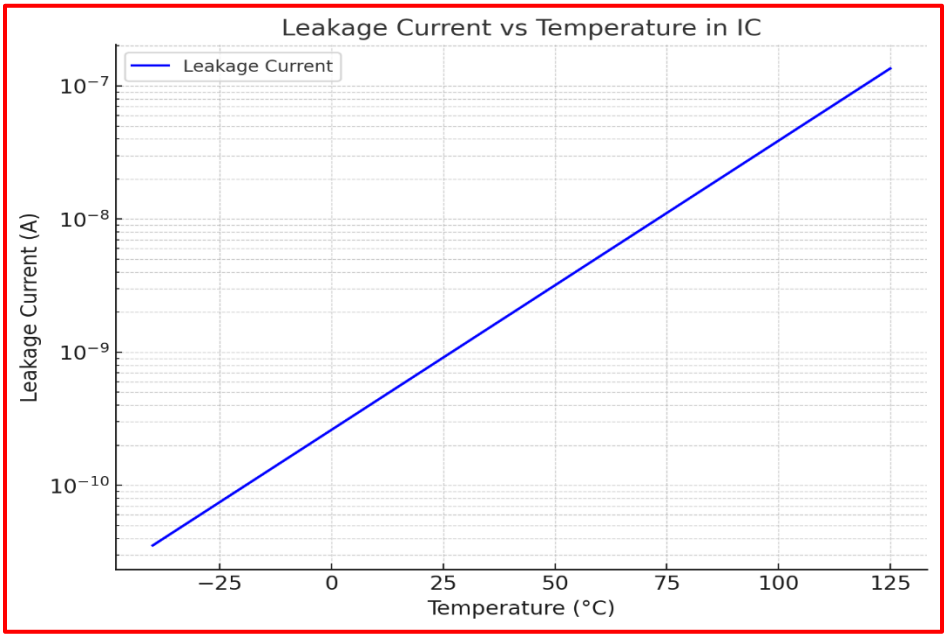


Figure 20. Leakage increase as a function of temperature

Feedback Loop of Temperature and Leakage:

- Self-Reinforcing Loop:
  - Increased dynamic power causes a temperature rise.
  - Higher temperatures increase leakage currents.
  - Increased leakage power further raises the temperature, exacerbating the problem.
  - This loop can potentially lead to thermal runaway if not properly managed.

Graphs Illustrating the Relationships:

  - A graph showing the compounding effect of increasing temperature on total power consumption, illustrating the potential for thermal runaway if not controlled.

Conclusion

The increase in dynamic power is due to higher transistor count, higher operating frequency, and operating voltage squared value and this all leads to elevated temperatures in ICs. This rise in temperature, in turn, increases leakage currents, further compounding power consumption and temperature rise. Effective thermal management and power optimization strategies are crucial to prevent thermal runaway and ensure the reliable operation of modern ICs.

Thermal Considerations and Management Techniques

Thermal Limits in ICs

Thermal limits are critical in the design and operation of integrated circuits. Managing heat effectively is essential to ensure the reliability, performance, and longevity of ICs. Here's an overview of the thermal limits and considerations for ICs:

Thermal Limits of Integrated Circuits

- **Junction Temperature (Tj)**:
  - **Definition**—The temperature of the semiconductor junction where the transistor action occurs. It is the most critical temperature limit for ICs.
  - **Typical Limits**—ICs typically operate safely up to junction temperatures of around 85°C to 125°C. Some high-performance ICs can handle even higher temperatures, up to 150°C, but this depends on the design and materials used.
- **Ambient Temperature (Ta)**:
  - **Definition**—The temperature of the environment surrounding the IC. This affects the overall thermal management of the IC package.
  - **Typical Limits**—Standard operating ambient temperatures are usually in the range of 0°C to 70°C for commercial applications and -40°C to 85°C for industrial-grade components.
- **Thermal Resistance (Rθ)**:
  - **Definition**—The measure of a package's ability to conduct heat away from the junction to the surrounding environment. It's expressed as °C/W (degrees Celsius per watt).
  - **Components**—Includes junction-to-case thermal resistance (Rθjc) and junction-to-ambient thermal resistance (Rθja).
- **Heat Dissipation**:
  - **Definition**—The process of transferring heat away from the IC to prevent overheating. Effective heat dissipation involves using heat sinks, thermal pads, or advanced packaging techniques.
  - **Methods**—Methods include using thermal vias, improving airflow with fans, or integrating heat spreaders and sinks.

- **Thermal Runaway**:
  - **Definition**—A situation where an increase in temperature leads to further increases in power dissipation, potentially causing the IC to fail.
  - **Prevention**—Effective thermal management and designing circuits with built-in thermal protection mechanisms can help avoid thermal runaway.

Thermal Management Techniques

- **Heat Sinks and Spreaders**:
  - Attach to the IC package to increase surface area and enhance heat dissipation.
- **Thermal Pads and Grease**:
  - Used between the IC and heat sink to improve thermal conductivity and reduce thermal resistance.
- **Thermal Vias**:
  - Placed in PCBs to conduct heat from the IC to other layers of the board or heat sinks.
- **Advanced Packaging**:
  - Technologies like CoWoS (Chip-on-Wafer-on-Substrate) and 3D ICs incorporate thermal management features into the package design.
- **Fan and Airflow Management**:
  - Utilize fans and proper airflow design to help dissipate heat from the IC and surrounding components.
- **Liquid Cooling**:
  - In high-performance or high-density applications, liquid cooling systems can be used to manage thermal loads effectively.

Design Considerations

- **Power Consumption**—Higher power consumption generates more heat, so optimizing power usage is crucial for thermal management.
- **Operational Environment**—Consider the ambient temperature and cooling capabilities available when designing and selecting ICs.
- **Reliability and Longevity**—Continuous operation at high temperatures can degrade the IC's performance and lifespan, so maintaining temperatures within safe limits is essential.

Effective thermal management is key to ensuring that integrated circuits operate reliably and efficiently, avoiding thermal-induced failures and maintaining optimal performance.

Thermal Management Details

Thermal management in electronic packaging is essential for maintaining the performance and reliability of integrated circuits and other electronic components. Effective thermal solutions help dissipate heat generated during operation and prevent overheating. Table 21 shows some common and advanced package thermal solutions:

Table 21. Advanced Packaging Thermal Solutions

Solution	Description
Heat Sinks	• Description: Heat sinks are metallic components (often made of aluminum or copper) attached to the IC package to increase the surface area available for heat dissipation. • Types: – Pin-Fin Heat Sinks: Feature multiple pins or fins to increase the

Solution	Description
	surface area and improve heat dissipation. – Heat Spreaders: Spread heat more evenly across the surface to prevent hot spots.
Thermal Pads and Grease	• Thermal Pads: – Description: Soft, conformable materials placed between the IC and heat sink or enclosure to improve thermal contact. – Benefits: Help fill air gaps and improve thermal conductivity. • Thermal Grease/Paste: – Description: A viscous material applied between the IC and heat sink to enhance thermal transfer. – Benefits: Fills microscopic gaps and improves thermal interface conductivity.
Thermal Vias	• Description: Holes in the PCB (printed circuit board) filled with conductive material (e.g., copper) to transfer heat from the IC through the PCB layers to a heat spreader or heat sink. • Types: – Blind Vias: Extend only partway through the PCB. – Through-Hole Vias: Extend completely through the PCB.
Thermal Interface Materials (TIMs)	• Description: Materials used between the IC and heat sink to improve thermal conductivity. They include thermal pads, grease, and phase-change materials. • Types: – Phase-Change Materials: Change from solid to liquid at operating temperatures to improve thermal contact. – Thermal Gels: Provide a flexible interface with high thermal conductivity.
Advanced Packaging Technologies	• Heat Spreaders: – Description: Metal plates integrated into the IC package to distribute heat more evenly and improve heat dissipation. • Integrated Heat Sinks: – Description: Heat sinks that are part of the package design, often including features like fins or embedded heat pipes. • Chip-on-Wafer-on-Substrate: – Description: A packaging technology that uses a silicon interposer to connect multiple chips, improving thermal and electrical performance. • 3D ICs: – Description: Stacked ICs with through-silicon vias to connect multiple layers, which requires advanced thermal management to handle heat generated by densely packed components.
Liquid Cooling	• Description: Uses liquids (often water or specialized cooling fluids) to absorb and transport heat away from the IC. • Types: – Direct Liquid Cooling: Liquid is directly in contact with the IC or heat spreader. – Cold Plates: Metal plates with embedded channels through which coolant flows.
Forced Air Cooling	• Description: Uses fans to direct airflow over the IC or heat sink to enhance heat dissipation. • Types: – Axial Fans: Move air parallel to the fan's axis. – Blower Fans: Move air perpendicular to the fan's axis, often used in confined spaces.
Heat Pipes	• Description: Devices that transfer heat via phase change and capillary action within a sealed tube. • Applications: Often used in conjunction with heat sinks or heat spreaders to enhance heat transfer.

Solution	Description
Thermal Management Coatings	• Description: Special coatings applied to the IC or PCB to enhance thermal conductivity or provide thermal insulation. • Types: – Thermal Conductive Coatings: Improve heat transfer between components. – Thermal Barrier Coatings: Provide thermal insulation to protect sensitive areas.
Design Considerations	• Thermal Simulation: Use of software tools to model and analyze heat distribution and manage thermal performance during the design phase. • Material Selection: Choosing materials with appropriate thermal properties (e.g., high thermal conductivity for heat spreaders).

Effective thermal management solutions are crucial for ensuring the optimal performance, reliability, and longevity of electronic devices. The choice of solution depends on the specific application, power dissipation requirements, and available space. It is therefore necessary to look at all these elements in greater depth.

Force Air and Water Cooling

Forced Air and Water Cooling are two common methods used to manage the thermal output of IC packages, each with its own set of limitations and considerations. **Forced air cooling** involves using fans or blowers to move air over the IC package, enhancing heat dissipation through convection. **Water cooling** involves using water or a water-glycol mixture to transfer heat away from the IC package. This method is often employed in high-performance or high-power applications.

Table 22 compares forced air and water cooling advantages, limits, and challenges associated with each cooling method:

Table 22. Comparison of Forced Air and Water Cooling

Type of Cooling	Advantages	Limits and Challenges
Forced Air Cooling	• Cost-Effective: Fans and heat sinks are relatively inexpensive and easy to implement. • Simplicity: Air cooling systems are simpler to design and install compared to liquid cooling systems. • Maintenance: Fans are generally easy to maintain and replace.	**Heat Dissipation Limits** • Limited Capacity: Air cooling is effective for moderate heat loads but can struggle with very high thermal outputs. As the heat generated by the IC increases, the effectiveness of air cooling diminishes, leading to potential thermal throttling or overheating. • Thermal Resistance: The thermal resistance of air is relatively high compared to liquids, limiting how quickly heat can be dissipated. **Airflow Management** • Non-Uniform Cooling: Achieving uniform airflow over all parts of the IC package can be challenging. Hot spots may still form if airflow is not evenly distributed. • Dust and Contamination: Dust accumulation on fans and heat sinks can reduce cooling efficiency over time and require periodic cleaning. **Noise** • Fan Noise: Fans can generate significant noise, which might be undesirable in noise-sensitive environments. **Size and Space Constraints** • Bulkiness: The size of the heat sink and fan assembly can be substantial, potentially limiting design options and increasing the overall size of the package.

Type of Cooling	Advantages	Limits and Challenges
Water Cooling	• High Heat Capacity: Water has a much higher thermal conductivity and heat capacity than air, allowing for more efficient heat transfer and dissipation. • Compact Design: Water cooling systems can be more compact compared to air cooling systems because they can transfer a larger amount of heat with a smaller physical footprint.	**Complexity and Cost** • System Complexity: Water cooling systems are more complex to design, install, and maintain. They require pumps, radiators, tubing, and often a reservoir. • Higher Cost: The components and installation of a water cooling system are generally more expensive compared to air cooling solutions. **Leakage Risk** • Potential Leaks: Water cooling systems carry the risk of leaks, which can damage electronic components and lead to failure. Proper sealing and regular maintenance are essential to mitigate this risk. **Maintenance** • Regular Maintenance: Water cooling systems require regular maintenance to prevent issues such as algae growth, corrosion, or sediment build-up. The coolant needs to be replaced periodically. **Cooling Efficiency** • Dependence on Radiator Size: The efficiency of a water cooling system depends heavily on the size and effectiveness of the radiator. If the radiator is not sufficiently large or well-placed, cooling performance can be compromised. **Weight and Space** • System Weight: While the cooling components are often compact, the overall system (including the water tank and tubing) can be heavy and require additional space.

Figure 21. Example of air cooling of semiconductor boards

Summary

Forced Air Cooling:

- **Limits—**Less effective for very high heat loads, potential for uneven cooling, noise, and bulkiness.
- **Best For—**Moderate heat dissipation needs, cost-sensitive applications, and where space allows for the necessary cooling components.

Water Cooling:

- **Limits—**Higher complexity and cost, risk of leakage, maintenance requirements, and dependency on radiator efficiency.
- **Best For—**High-performance or high-power applications where compact design and superior heat dissipation are critical.

Both cooling methods have their own strengths and weaknesses, and the choice between them depends on factors such as the power dissipation requirements, space constraints, cost considerations, and the specific application.

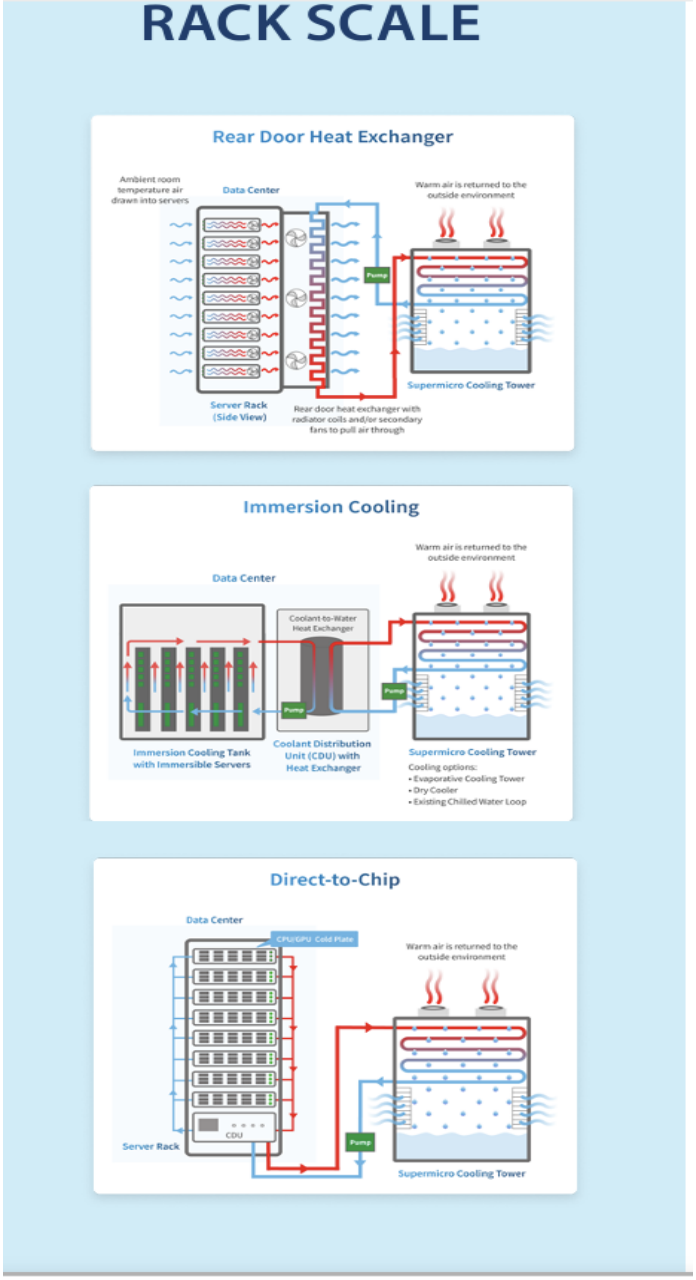


Figure 22. Schematic example of water cooling of boards in a AI rack

Water Cooling Details

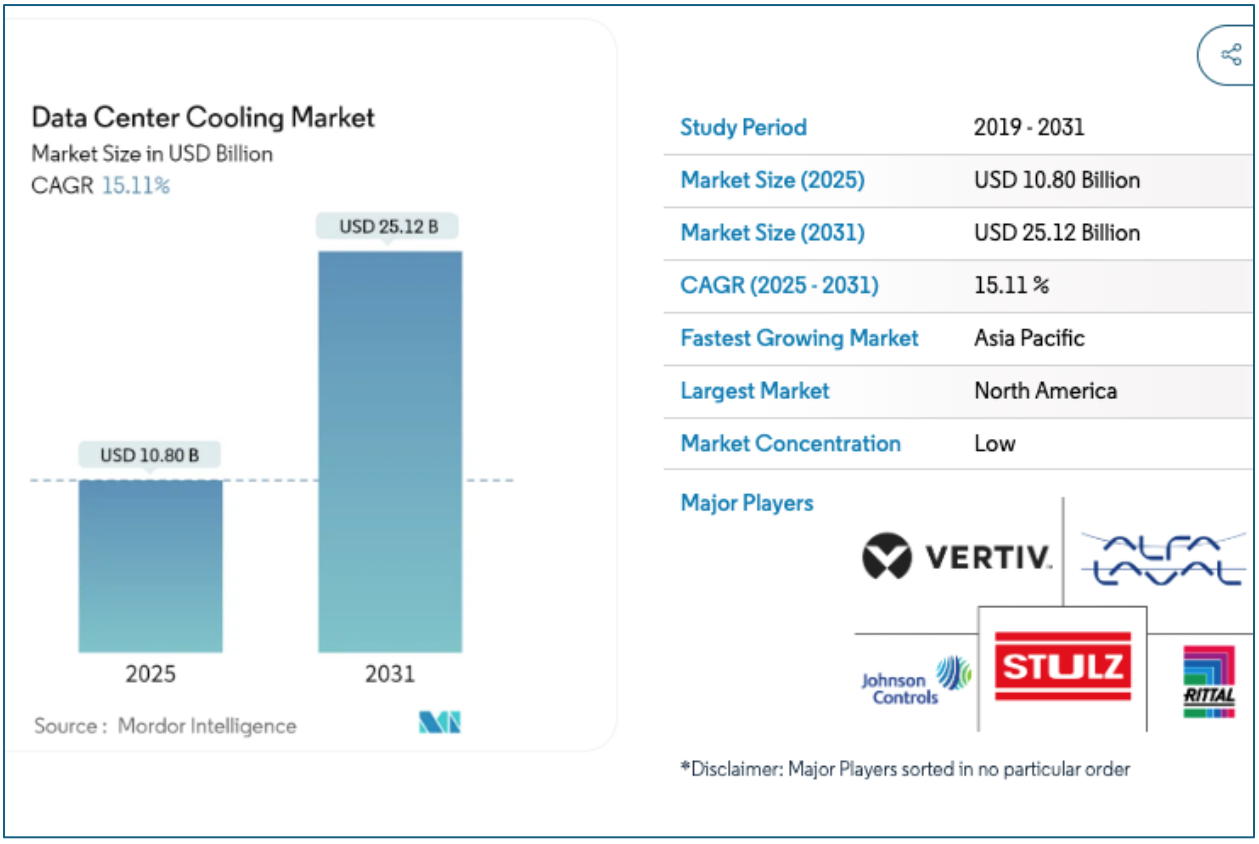
IC (integrated circuit) encapsulation for water cooling is a method used to manage the heat generated by high-performance electronics. Effective thermal management is crucial in applications where ICs operate at high power levels or in compact, high-density environments. Water cooling is one of the advanced techniques employed to achieve efficient heat dissipation. Here's a detailed overview of the IC encapsulation process tailored for water cooling:

Table 23. IC Encapsulation Process Tailored for Water Cooling

Process	Description
Design Considerations	**Thermal Analysis** • Heat Generation: Assess the heat generated by the IC and the thermal requirements for cooling. • Thermal Resistance: Calculate the thermal resistance of the package to determine how effectively heat can be transferred away from the IC. **Water Cooling Integration** • Cooling System Design: Design the water cooling system, including the coolant flow paths, heat exchangers, and pumps, to ensure effective heat removal.
IC Encapsulation	**Substrate Selection** • Material Choice: Select substrates with good thermal conductivity, such as copper or ceramic, that can facilitate efficient heat transfer from the IC to the cooling system. **Die Attachment** • Thermal Interface Material (TIM): Apply a TIM (such as thermal grease, thermal pads, or phase-change materials) between the IC die and the substrate to enhance thermal conductivity. • Die Bonding: Attach the IC die to the substrate using an adhesive that can handle thermal expansion and high temperatures. **Encapsulation Material** • Protective Coating: Use a high-thermal-conductivity encapsulation material (e.g., epoxy or silicone) that provides protection while allowing efficient heat transfer. Ensure that the encapsulation material has properties compatible with water cooling. **Water Cooling Integration** • Cooling Plates: Attach or integrate a water-cooled heat sink or cooling plate directly to the encapsulated IC. These plates have channels or fins designed to facilitate water flow and enhance heat dissipation. • Sealing and Gaskets: Use seals or gaskets to prevent water leakage and ensure that the cooling system is tightly integrated with the encapsulated IC.
Assembly and Integration	**Heat Spreader or Heat Sink** • Attachment: Attach a heat spreader or heat sink to the encapsulated IC. This component is often in direct contact with the cooling plates and helps to distribute heat evenly. • Water Blocks: In more advanced designs, use custom water blocks that are designed to fit precisely over the IC and channel coolant directly over the heat-generating areas. **Cooling System Connection** • Fluid Routing: Connect the cooling system's fluid channels to the water blocks or heat spreaders. Ensure that the connections are secure and that there are no leaks. • Coolant Flow: Design the coolant flow to maximize contact with the heat spreader or cooling plate. Proper flow design ensures efficient heat removal.
Testing and Validation	**Leak Testing** • Pressure Testing: Conduct pressure testing to ensure that there are no leaks in the water-cooling system and that the encapsulation is securely sealed. • Leak Detection: Use leak detection methods to confirm the integrity of the system. **Thermal Performance Testing** • Temperature Measurements: Measure the temperature of the IC and the heat spreader to verify that the water cooling system is effectively removing heat. • Cooling Efficiency: Evaluate the efficiency of the water cooling system in maintaining the IC within its operational temperature range. **Reliability Testing** • Long-Term Testing: Perform long-term reliability testing to ensure that the encapsulation and cooling system can withstand extended use and thermal cycling.
Final Assembly	**Encapsulation of Final Product**

Process	Description
	• Protective Housing: Place the encapsulated IC and cooling system into a protective housing or enclosure, if necessary, to provide additional protection and integration into the final product. **Integration into Systems** • System Integration: Integrate the cooled IC into the larger electronic system or device, ensuring that all connections and interfaces are properly aligned and functional.
Maintenance and Upkeep	**Regular Checks** • Coolant Levels: Monitor coolant levels and replace or top up as needed. • System Integrity: Regularly check for any signs of leakage or wear in the cooling system. **Cleaning and Servicing** • Cooling System Cleaning: Clean the cooling system periodically to remove any debris or contaminants that may affect performance.

Using water cooling for ICs provides superior thermal management compared to air cooling, especially in high-power and high-density applications. However, it requires careful design and integration to ensure reliability and effectiveness.



Source: Mordor Intelligence

Figure 23. Projected data center cooling market size

8th ETP Summary—Thermal Limits and Management

Thermal limits are critical in the design and operation of integrated circuits. Managing heat effectively is essential to ensure the reliability, performance, and longevity of ICs. The increase in dynamic power due to higher transistor count, operating frequency, and voltage leads to elevated temperatures in ICs. This rise in temperature, in turn, increases leakage currents, further compounding power consumption and temperature rise. Effective thermal management and power optimization strategies are crucial to prevent thermal runaway and ensure the reliable operation of modern ICs. Most of all as power dissipation increases in any IC it provokes temperature to raise causing leakage current to increase until functionality of the transistor is severely affected and eventually the transistor will cease to properly function. **Forced Air and Water Cooling** are two common methods used to manage the thermal output of IC packages, each with its own set of limitations and considerations. The choice between them depends on factors such as the power dissipation requirements, space constraints, cost considerations, and the specific application.

Concluding Remarks

Encapsulation of transistors was devised early in IC design and manufacturing to prevent any damage to the frail structure of transistors during handling. The package represented also the means of connecting the input and output electrical signals generated by the IC via the package to the other components located on the same PCB board. As complexity of ICs kept on increasing it became necessary to evolve the package to a dual in-line structure to accommodate for pins on two sides. Eventually the package evolved to a square format with pins on all four sides to maximize the number of available inputs/outputs; eventually the IC became the so called “flip chip” connecting itself directly to metal traces located on the board. Nevertheless, the role of packaging remained limited to physical protection of ICs and as the electrical conduit to the board.

However, things drastically changed in the past 15 years as the role of packaging surged to the level of an active system solution provider as progressively it became possible placing multiple dice in the same package eliminating the need to individually package each die and with substantially improvements in signal propagation delay due to the reduced interconnect lengths. The growth in popularity of using the data center and the ubiquitous accessibility to any form of data via wired and wireless communications has made possible to increase the power level handled by stationary ICs from the one hundred of watts of the PC era to Kilowatts level without melting the IC by removing the generated heat by means of air and water cooling.

The EPT Parts 2 and 3 detail the technologies associated with advanced packaging and future prospects with high-performance computing and artificial intelligence.

Glossary/Abbreviations

Term	Definition
3DIC	3 dimensional IC
ABF	Ajinomoto build-up film
ACF	Anisotropic conductive films
ADAS	Advanced driver-assisted system
AI	Artificial intelligence
AiP	Antenna in package
AlN	Aluminum nitride
ASIC	Application specific IC
BC	Backside clocking
BeO	Beryllium Oxide
BGA	Ball grid array
BGA-HS	Ball grid array with heat spreader
BiCMOS	Bipolar CMOS
BOM	Bill of materials
BPD	Backside power distribution
BT	Bismaleimide triazine
C4	Controlled Collapse Chip Connection
CD	Clock distribution
CDN	Clock distribution network
CMOS	Complimentary metal-oxide semiconductor
CMP	Chemical-mechanical polishing
COB	Chip on board
CoWoS	Chip-on-wafer-on-substrate
CPU	Central processing unit
CSBGA	Chip-scale BGA
CSP	Chip-scale package
DCRI	Data center rack integration
DIP	Dual inline package
DRAM	Dynamic random access memory
DR-BGA	Dual-row BGA
DSP	Digital signal processor
E/O	Electrical to optical
EMI	Electromagnetic interference
EMIB	Embedded multi-die interconnect bridge
EPT	Executive Packaging Tutorial
EUV	Extreme ultraviolet
eWLB	Embedded wafer-level ball grid array
FBGA	Fine pitch BGA
FC	Flip chip
FET	Field-effect transistor
FOWLP	Fan-out wafer-level packaging
FPGA	Field-programmable gate array
FR-4	Flame retardant-4
GB	Gigabyte
GIS	Global integrated system
GPU	Graphic processing unit
HB	Hybrid bonding
HBM	High bandwidth memory
HDI	High-density interconnect
HPC	High-performance computing
HVM	High volume manufacturing
I/O	Input/output

IBM	International Business Machines
IC	Integrated circuit
IEEE	Institute of Electrical and Electronics Engineers
IFT	International Focus Team
InFO	Integrated Fan Out
IoT	Internet of Things
IRDS	International Roadmap for Devices and Systems
ITRS	International Technology Roadmap for Semiconductors
LED	Light emitting diode
LGA	Land grid array
LLM	Large language model
LPDDR	Low power double data rate
MCM	Multi-chip modules
MEMS	Microelectromechanical systems
ML	Machine learning
MOF	Metal-organic frameworks
MPU	Multiprocessor unit
NIC	Network interface card
NOC	Network on chip
O/E	Optical to electrical
PC	Personal computer
PCB	Printed circuit board
PCM	Phase-change material
PDN	Power delivery network
POP	Package on package
QFN	Quad flat no lead
QFP	Quad flat package
R0	Thermal resistance
RDL	Redistribution layer
RF	Radio frequency
RoHS	Restriction of Hazardous Substances
SAC	Sn/Ag/Cu
SBGA	Standard BGA
SiB	System in board
SiC	Silicon carbide
SiP	System in package
SMD	Surface mount device
SMT	Surface mount technology
SoC	System on chip
SPD	Signal propagation delay
SRAM	Static random access memory
SSD	Solid state drive
STA	Static timing analysis
STCO	System technology co-optimization
Ta	Ambient temperature
TB	Terabyte
TGV	Through glass via
TIA	Transimpedance amplifier
TIM	Thermal interface materials
Tj	Junction temperature
TSMC	Taiwan Semiconductor Manufacturing Company
TSV	Through silicon via
TWG	Technology Working Group
UV	Ultra violet

Appendices

Appendix A—Substrate Packaging Sizes and the Substrate Fabrication Process

Substrate Packaging Sizes

Table A1. Typical sizes for different types of packaging substrates

Substrate Type	Packaging	Size Range
Printed Circuit Board Substrates	For standard packages like BGA, QFN, etc.	Typically, between 10 mm x 10 mm to 50 mm x 50 mm, depending on the complexity and number of connections.
	For larger systems or modules:	Can go up to several hundred millimeters on each side, such as 150 mm x 150 mm or even larger for complex systems like servers or high-performance computing modules.
Wafer-Level Packaging (WLP) Substrates	For Fan-Out Wafer-Level Packaging (FOWLP)	Substrates can be as large as the wafer itself, typically ranging from 200 mm to 300 mm in diameter. Individual die on these wafers might range from 1 mm^2 to 20 mm^2.
Ceramic Substrates	For high-reliability applications (e.g., aerospace, automotive):	Commonly range from 10 mm x 10 mm to 50 mm x 50 mm, similar to organic substrates, though they can be larger depending on the application.
	For power electronics modules:	These might be much larger, up to 100 mm x 100 mm or more.
System-in-Package and 2.5D/3D IC Substrates	SiP	Substrates can range from 10 mm x 10 mm to around 30 mm x 30 mm, depending on how many components are integrated.
	2.5D/3D IC Substrates	Typically, these substrates are slightly larger than the die themselves, ranging from 20 mm x 20 mm to 40 mm x 40 mm or more.
Custom and Advanced Substrates	Custom or advanced applications (e.g., AI processors, high-performance computing)	Can be quite large, sometimes exceeding 100 mm on a side, especially in cases where multiple large die are integrated on a single substrate.

Manufacturing Process of Substrate Packaging—Detailed Steps

Table A2. Substrate Packaging Process Steps

Process Name	Description
1. Substrate Material Preparation	Material Selection: Choose the appropriate substrate material, such as FR-4 (a type of fiberglass), BT (Bismaleimide Triazine), or ceramic, depending on the application's thermal and electrical requirements. Lamination: Multiple layers of dielectric material are laminated together, typically using prepregs (pre-impregnated materials) to form a multi-layer substrate. This step ensures the substrate has the necessary mechanical strength and electrical isolation between layers.
2. Circuit Patterning	Photolithography: Apply a photoresist layer on the substrate's surface, then expose it to ultraviolet (UV) light through a photomask that defines the desired circuit pattern. The exposed areas are chemically developed, revealing the underlying material. Etching: The exposed areas of the substrate are etched away using a chemical etchant, leaving behind the desired circuit patterns (e.g., traces, pads, and vias). Stripping: The remaining photoresist is stripped away, leaving only the etched circuit patterns on the substrate.
3. Via Formation	Drilling: Vias (vertical interconnect accesses) are drilled into the substrate to allow electrical connections between different layers. Depending on the design, laser drilling or mechanical drilling is used. Via Metallization: The drilled vias are coated with a conductive material, typically copper, using processes like electroless plating or sputtering. This step ensures electrical connectivity between layers.
4. Copper Plating	Surface Plating: A thin layer of copper is plated onto the surface of the substrate and within the vias to provide a conductive pathway for the circuit. Pattern Plating: Additional copper is selectively plated onto the circuit patterns to achieve the desired thickness and conductivity.
5. Surface Finishing	Solder Mask Application: A solder mask (a protective coating) is applied over the substrate's surface, leaving openings only where soldering will occur. This step prevents solder bridges and protects the substrate from oxidation and contamination. Surface Finish: Various surface finishes, such as ENIG (Electroless Nickel Immersion Gold), OSP (Organic Solderability Preservative), or HASL (Hot Air Solder Leveling), are applied to the exposed copper areas. These finishes ensure good solderability and protect the copper from oxidation.
6. Assembly Process	Die Attach: The IC (die) is attached to the substrate using an adhesive, typically an epoxy resin. The adhesive is cured to secure the die in place. Wire Bonding or Flip-Chip: Depending on the packaging type, wire bonding (connecting the die to the substrate using fine metal wires) or flip-chip (attaching the die directly to the substrate via solder bumps) is performed to establish electrical connections.

Process Name	Description
	Encapsulation: The assembled substrate and die are encapsulated with a protective material, such as an epoxy mold compound, to protect the IC from environmental factors and mechanical stress.
7. Testing and Inspection	Electrical Testing: The packaged substrate undergoes electrical testing to verify functionality, including continuity, isolation, and signal integrity. Visual Inspection: Automated optical inspection (AOI) systems are used to detect any defects, such as misalignment, solder bridges, or cracks. X-ray Inspection: X-ray inspection is often used to inspect hidden features, such as solder joints in flip-chip packages.
8. Final Finishing	Singulation: The packaged substrates are separated (singulated) from the panel, often using mechanical sawing, laser cutting, or other precision cutting methods. Marking and Labeling: The individual substrates are marked with identification codes, logos, or other necessary information using laser or inkjet marking.
9. Final Testing and Quality Control	Final Testing: The finished substrate packages undergo final testing to ensure they meet all electrical, mechanical, and thermal specifications. Quality Control: A final quality control check ensures that the packaging meets all required standards and specifications. This includes visual inspection, dimensional verification, and performance testing.
10. Packaging and Shipping	Packaging: The finished substrate packages are carefully packed in protective materials to prevent damage during shipping. Shipping: The packaged substrates are shipped to customers or assembly facilities for further integration into larger electronic systems.

Key Message

The manufacturing process flow for substrate packaging involves several precise steps, from material preparation to final testing. Each step is crucial in ensuring that the substrate package performs reliably in its intended application. Advances in substrate materials, process technologies, and quality control methods continue to drive improvements in performance, miniaturization, and cost-effectiveness in semiconductor packaging. Substrate size is continuously increasing as the SiP is absorbing more and more of the components located on the PCB.

Appendix B—Land Grid Array Packaging Fabrication Process

Table A3. LGA fabrication process

Process Name	Description
1. Package Design and Substrate Preparation:	Design: The package design is created using CAD tools, defining the layout of the pads, traces, and other elements. Substrate Fabrication: The substrate, typically made from organic materials like FR-4 or high-density interconnect (HDI) substrates, is prepared. It includes multiple layers of conductive traces and insulating materials.
2. Pad Patterning:	Photoresist Application: A photoresist layer is applied to the substrate. Photolithography: The design pattern is transferred onto the photoresist using UV light through a mask. Etching: The exposed areas are etched away to form the pad pattern on the substrate.
3. Via Formation:	Drilling: Holes (vias) are drilled through the substrate to connect different layers. Via Filling: The vias are filled with a conductive material, such as copper, to create electrical connections between layers.
4. Die Attach and Wire Bonding:	Die Attach: The IC die is attached to the substrate using die-attach adhesives or solder. Wire Bonding: Thin wires are bonded between the die's pads and the package's internal traces. For some LGA packages, flip-chip bonding may be used instead.
5. Solder Mask and Surface Finish:	Solder Mask: A solder mask is applied to protect the package from solder bridging and environmental damage. Surface Finish: Surface finishes, such as electroless nickel immersion gold (ENIG) or immersion silver, are applied to the pads to ensure good solderability and protection against oxidation.
6. Inspection and Testing:	Inspection: The package undergoes visual and automated optical inspection (AOI) to ensure that the pads and connections are correctly formed. Testing: Electrical testing is performed to check for connectivity and performance.
7. Final Assembly:	Mounting: The completed LGA package is mounted onto the PCB. The package is aligned with corresponding pads on the PCB. Soldering: The package is soldered onto the PCB using reflow soldering. Unlike BGAs, LGA packages do not use solder balls but rely on flat pads that are soldered directly.

Appendix C—Ball Grid Array Fabrication and Electrical Connection Processes

BGA Fabrication Process

Table A4. Ball Grid Array Fabrication Process

Process	Description
1. Die Preparation	• The IC die is prepared and tested before being packaged.
2. Package Substrate	• The BGA package consists of a substrate, usually made of epoxy resin or other materials, which provides the base for the solder balls and the IC die.
3. Solder Ball Placement	• Solder balls are placed on the underside of the package substrate. This is typically done using a stencil and a solder paste screen-printing process or by placing pre-formed solder balls.
4. Die Attachment	• The IC die is attached to the substrate using a die-attach adhesive or solder. Wire bonding or flip-chip technology may be used to connect the die to the package.
5.Encapsulation	• The die and its connections are encapsulated with a protective resin to shield them from environmental factors and mechanical damage.
6.Reflow Soldering	• The package is subjected to a reflow soldering process, where the solder balls are heated to melt and then solidify, forming a reliable connection between the package and the PCB.
7. Inspection and Testing	• The final package undergoes inspection and testing to ensure that it meets quality and performance standards.

BGA Electrical Connection Process

Table A5. BGA Electrical Connection Process

Process	Description
1. Die Preparation and Attachment	• Die Bonding: The IC die (the actual semiconductor component) is first attached to the package substrate. This is typically done using a die-attach adhesive or a solder-based die attach method. In some cases, flip-chip technology may be used, where the die is flipped and bonded directly to the package substrate with solder bumps.
2. Wire Bonding or Flip-Chip Bonding	• Wire Bonding: In traditional BGA packages, tiny gold or aluminum wires are used to connect the die's pads to the package substrate's leads or pads. The wires are bonded using a wire bonding machine that performs ultrasonic or Thermosonic bonding. These wires create the electrical path between the die and the package. • Flip-Chip Bonding: In flip-chip BGA packages, the die is equipped with solder bumps on its surface, which are aligned with corresponding pads on the package substrate. These solder bumps are reflowed to create electrical connections directly between the die and the substrate, eliminating the need for wire bonding.
3. Package Substrate	• Substrate: The package substrate acts as an intermediary layer between the die and the PCB. It usually consists of multiple layers of insulating and conductive materials that form the internal routing for electrical signals. The substrate includes pads or traces connected to the solder balls on the package's underside.
4. Solder Ball Attachment	• Solder Balls: Solder balls are placed in an array pattern on the underside of the BGA package. These balls are typically made of a tin-lead or lead-free solder alloy. During the assembly process, the solder balls are attached to the package substrate using solder paste and reflowed to form solid connections.
5. Reflow Soldering	• Reflow Process: The BGA package, with the solder balls in place, undergoes a reflow soldering process. In this process, the entire package is heated in a reflow oven, causing the solder balls to melt and form a permanent connection between the package substrate and the PCB. Upon cooling, the solder solidifies, creating a reliable electrical and mechanical bond.
6. Final Assembly and Inspection	• Mounting on PCB: The BGA package is then placed onto the PCB in the correct alignment with corresponding PCB pads. • Inspection: After mounting, the connections are inspected to ensure proper alignment and soldering. X-ray inspection is often used to check the integrity of the solder joints, as the solder balls are not visible from the top.

Appendix D—Flip Chip Packaging Fabrication Process

Table A6. Flip Chip Packaging Process

Process	Description
Chip Preparation:	The semiconductor chip is fabricated with bumps of solder (often referred to as "solder bumps" or "bumps") on its surface. These bumps serve as both the electrical connections and the mechanical support for the chip.
Flipping the Chip:	The chip is flipped so that the side with the solder bumps faces downwards, towards the package substrate or another chip.
Attachment	The solder bumps are reflowed (heated until they melt and then solidify) to create electrical connections between the chip and the package substrate. This process is called "solder bumping" or "flip chip bonding."
Encapsulation	The chip is encapsulated or covered with a protective material to shield it from physical damage and environmental factors.
Final Testing:	The packaged chip undergoes final testing to ensure that all connections are functioning correctly and that the chip meets performance specifications.

Appendix E—Flip Chip Manufacturing Process

The flip chip manufacturing process is a popular method used in semiconductor packaging, where the chip is flipped upside down and bonded directly to the package substrate. This approach offers several advantages, including improved electrical performance, reduced package size, and enhanced thermal dissipation. Here's a detailed description of the flip chip manufacturing process:

Table A7. Flip Chip Manufacturing Process

Process	Steps	Description
Wafer Preparation	**Wafer Fabrication**	• Semiconductor Wafer: The process begins with a semiconductor wafer, typically made of silicon, on which integrated circuits are fabricated. • Front-End Processing: This involves photolithography, etching, doping, and other processes to create the ICs on the wafer.
	Wafer Testing	• Electrical Testing: Before proceeding, each wafer is tested for functionality and performance to ensure that the ICs meet specifications. This step helps identify any defective chips.
Underfill and Bumping	**Bumping**	• Formation of Bumps: Small metal bumps (typically made of solder) are deposited on the wafer's surface. These bumps serve as the electrical and mechanical connections between the chip and the substrate. Common methods for bump formation include solder ball placement, electroplating, or screen printing. • Reflow Soldering: The wafer is heated in a reflow oven to melt the solder bumps, which then solidify as the wafer cools, forming well-defined solder bumps.
	Underfill Preparation	• Underfill Material: An underfill material, often an epoxy-based resin, is applied to the area around the bumps to fill the gap between the chip and the substrate after the chip is flipped and attached. The underfill material improves mechanical strength and thermal performance.
Flip Chip Assembly	**Flip Chip Bonding**	• Flip the Chip: The wafer is flipped so that the solder bumps face downward. The chip is then aligned with the package substrate or a carrier, which has corresponding pads or bumps for the chip's solder bumps. • Bonding: The chip is pressed down onto the substrate, and the solder bumps are reflowed again to form reliable electrical and mechanical connections between the chip and the substrate.
	Underfill Curing	• Dispense Underfill: Once the chip is bonded, underfill material is dispensed around the edges of the chip to fill the gap between the chip and the substrate. • Curing: The underfill is then cured, typically through heat or ultraviolet light, to harden and secure the chip in place.
Post-Bonding Processes	**Dicing**	• Wafer Dicing: After successful bonding and curing, the wafer is diced into individual chips (known as die) using a dicing saw or laser. This step separates the chips from each other.
	Packaging	• Encapsulation: The individual die may be encapsulated in a protective material, such as plastic or ceramic, to protect the chip and its connections. • Final Testing: Each chip is tested again to ensure functionality and performance after the flip chip assembly process.
	Soldering	• Final Solder Attach: For some packages, additional solder balls or pads are added to the substrate or package for mounting the IC onto a printed circuit board. These solder balls are typically attached using a reflow process similar to the bumping step.
Quality Control and Reliability Testing	**Visual Inspection**	• Inspection: The final package is inspected for any defects or irregularities in the solder bumps, underfill, and overall assembly.
	Reliability Testing	• Thermal Cycling and Stress Testing: The assembled IC is subjected to various reliability tests, including thermal cycling, mechanical stress, and humidity exposure, to ensure long-term durability and performance.

Appendix F—MCM Assembly Process

The following table show the typical sequence for the MCM assembly process.

Table A8. MCM Assembly Process

Process	Steps	Description
Design and Preparation	**Design Specification**	• Requirement Analysis: Define the functional and performance requirements for the multichip module, including electrical, thermal, and mechanical specifications. • Chip Selection: Choose the semiconductor chips that will be integrated based on their functionality and compatibility.
	Substrate Preparation	• Substrate Selection: Choose an appropriate substrate (e.g., ceramic, organic, or silicon) that will support the chips and provide necessary interconnections. • Cleaning and Conditioning: Clean the substrate to remove contaminants and apply any required surface treatments to enhance adhesion and bonding.
Chip Preparation	**Wafer Dicing**	• Dicing: Cut the semiconductor wafer into individual die (chips) using a dicing saw or laser. Each die is separated from the wafer to be used in the assembly.
	Die Attachment	• Die Bonding: Attach each chip to the substrate using an adhesive or solder. Techniques include epoxy die attach, solder bonding, or thermocompression bonding.
Electrical and Mechanical Interconnection	**Wire Bonding**	• Wire Bonding: Connect the chip's bonding pads to the substrate's interconnects using fine wire bonds. This process uses gold or aluminum wires and is performed using ultrasonic or Thermosonic bonding techniques.
	Flip-Chip Bonding	• Flip-Chip Assembly: Involves mounting chips onto the substrate with solder bumps (or balls) facing downward. The solder is reflowed to make electrical connections between the chip and substrate.
	Micropillar/Through-Silicon Via Integration	• TSV: For 3D packaging, TSVs are used to create vertical electrical connections between chips stacked on top of each other.
Encapsulation and Protection	**Encapsulation**	• Mold Encapsulation: The assembled chips and substrate are encapsulated in a protective material, such as epoxy or silicone, to shield against physical damage, moisture, and contaminants.
	Underfill	• Underfill Application: For flip-chip assemblies, an underfill material is dispensed around the solder bumps to enhance mechanical strength and thermal performance.
Testing and Quality Assurance	**Electrical Testing**	• Functional Testing: Verify the electrical performance and functionality of the multichip assembly. This includes checking for correct signal operation and power delivery.
	Thermal Testing	• Thermal Performance: Assess the thermal characteristics to ensure that heat dissipation is adequate and that the assembly operates within safe temperature ranges.
	Mechanical Testing	• Mechanical Reliability: Test the mechanical robustness of the assembly, including resistance to mechanical stresses and vibrations.
Final Processing	**Marking and Labeling**	• Identification: Apply identifying marks or labels to the final product for traceability and assembly identification.
	Inspection and Rework	• Visual and Automated Inspection: Conduct visual and automated inspections to detect any defects or inconsistencies. Perform any necessary rework or repairs.
Packaging	**Final Packaging**	• Package the completed multichip module for shipment or integration into larger systems. This may involve additional protective packaging to prevent damage during transport.

Appendix G—Substrate Preparation Process for Multi-Chip Assembly

Table A9. MCM Substrate Preparation Process for Multi-Chip Assembly

Process	Steps	Description
Substrate Fabrication	**Design and Layout**	• The substrate design includes routing for electrical signals, power distribution, and thermal management. CAD tools are used to create detailed designs and layouts.
	Layer Formation	• For multi-layer substrates, layers of copper and dielectric materials are stacked and laminated together. The copper layers form the conductive traces, while the dielectric layers provide insulation.
	Patterning	• Photolithography is used to pattern the conductive traces on the substrate. This involves applying a photoresist layer, exposing it to UV light through a mask, and etching away the unwanted material.
Via Formation	**Drilling or Laser Ablation**	• Holes or vias are created to connect different layers of the substrate. This can be done using mechanical drilling or laser ablation.
	Via Filling	• The vias are filled with a conductive material, such as copper, through electroplating or other deposition methods.
Solder Mask and Surface Finish	**Solder Mask Application**	• A solder mask is applied to prevent solder from bridging between traces and to protect the substrate from environmental damage.
	Surface Finish	• Surface finishes, such as immersion gold, electroless nickel, or tin-lead, are applied to the exposed copper pads to ensure good solderability and to protect against oxidation.
Chip Mounting and Bonding	**Die Attach**	• The ICs are attached to the substrate using die-attach adhesives or solder bumps in flip-chip configurations.
	Wire Bonding	• For wire-bonded MCMs, fine wires are bonded from the die pads to the substrate's bond pads.
	Flip-Chip Bonding	• For flip-chip configurations, solder bumps on the IC are reflowed to make direct connections with the substrate.
Encapsulation and Protection	**Encapsulation**	• The assembled MCM is encapsulated with a protective material to shield the ICs and connections from physical damage and environmental factors.
	Thermal Management	• Thermal pads or heat spreaders may be added to enhance heat dissipation from the ICs.
Testing and Inspection	**Electrical Testing**	• The completed MCM is subjected to electrical testing to ensure that all connections are functioning correctly.
	Visual Inspection	• Automated optical inspection (AOI) and X-ray inspection are used to check for defects and ensure proper soldering.

Appendix H—Micropillars Fabrication Process

The fabrication of micropillars for packaging involves a series of precise and sophisticated processes. Here is a typical sequence for fabricating micropillars.

Table A10. Micropillars Fabrication Process

Process	Steps	Description
Substrate Preparation	**Cleaning**	• The substrate (such as a silicon wafer or PCB) is cleaned to remove contaminants and particles that could affect the adhesion and quality of the micropillars.
	Surface Activation	• Techniques like plasma treatment may be used to activate the surface of the substrate to improve adhesion.
Photolithography (for Patterning applications in SiP)	**Photoresist Application:**	• A photoresist layer is applied to the substrate. This is a light-sensitive material that will be used to define the pattern of the micropillars.
	Soft Baking:	• The photoresist is baked at a low temperature to remove solvents and improve adhesion.
	Exposure:	• The substrate with photoresist is exposed to ultraviolet (UV) light through a mask that has the desired micropillar pattern. The UV light changes the chemical properties of the photoresist in the exposed areas.
	Development:	• The exposed photoresist is developed, leaving behind a patterned photoresist layer that will serve as a mask for subsequent processing steps.
Micropillar Formation	**Deposition (for Initial Layer)**	• A thin layer of material, such as metal (copper, nickel) or polymer, is deposited onto the substrate. Techniques like sputtering or electroplating are commonly used for metal deposition.
	Etching	• The material is etched away from the areas not protected by the photoresist, forming the micropillars. Etching can be done using wet or dry (plasma) etching techniques, depending on the material used.
	Removal of Photoresist	• After etching, the remaining photoresist is stripped away, leaving the micropillars in place on the substrate.
Electroplating (if applicable)	**Seed Layer Deposition**	• A thin layer of metal (such as copper) may be deposited on the substrate to act as a seed layer for electroplating.
	Electroplating	• A thicker metal layer is electroplated onto the micropillars to increase their height or add additional features. Electroplating involves immersing the substrate in a metal salt solution and applying an electrical current to deposit metal onto the exposed areas.
Final Processing	**Planarization**	• The surface may be planarized to ensure that the micropillars are at the desired height and to remove any excess material or irregularities.
	Inspection	• The fabricated micropillars are inspected using microscopy techniques to ensure they meet design specifications and quality standards.
Integration into Packaging	**Assembly**	• The micropillar structures are integrated into the packaging system. This may involve bonding the substrate with micropillars to other components or layers in a package.
	Soldering/Bonding	• In some designs, the micropillars are used for soldering or bonding to other parts of the electronic assembly, such as semiconductor chips or substrates.
Testing and Quality Assurance	**Functional Testing**	• The final packaged device is tested to ensure that the micropillars perform as expected, providing the necessary electrical, thermal, and mechanical properties.
	Reliability Testing	• Additional tests are conducted to assess the reliability and durability of the micropillars and the overall package.

Appendix I—TSV Fabrication Process

Table A11. Fabrication of TSVs

Process Step	Description
Wafer Preparation	• The TSV fabrication process begins with the preparation of the silicon wafer. The wafer is typically thinned to a desired thickness before the TSVs are created.
Drilling or Etching	• TSVs are formed by drilling or etching holes through the silicon wafer. This process can be achieved using techniques such as deep reactive ion etching (DRIE) or laser drilling. The size and spacing of the TSVs are carefully controlled to ensure proper electrical performance and mechanical stability.
Lining and Filling	• After the holes are created, they are lined with a conductive material, typically through a process called chemical vapor deposition (CVD) or physical vapor deposition (PVD). Common lining materials include copper or tungsten. The TSVs are then filled with a conductive material to create the vertical interconnects. Copper is frequently used for filling due to its excellent electrical conductivity.
Planarization	• Once the TSVs are filled, the wafer is subjected to a planarization process, such as chemical-mechanical polishing (CMP). This process ensures that the surface of the wafer is smooth and level, which is crucial for subsequent packaging and bonding steps.
Passivation and Metallization	• The wafer is then coated with a passivation layer to protect the TSVs and prevent any chemical reactions. Metallization layers are also added to establish electrical connections between the TSVs and the external circuitry.

Appendix J—Practical examples of actual SiP packaging technologies

SiP (System In Package) is an advanced semiconductor packaging technology that represents a significant evolution in integrated circuit packaging, designed to improve performance, integration density, and thermal management. SiP integrates multiple chiplets or dice into a single package using advanced 3D packaging techniques. This allows combining different types of chips (such as logic, memory, and analog) in one package to create a more compact and efficient system. **SiP** typically uses a silicon interposer or a similar substrate layer to facilitate the high-density interconnections between the co-packaged chips. This approach reduces power consumption by minimizing the length of interconnects and improving signal integrity.

There are several packaging technologies, chip architectures and systems that provide solutions for building and implementing SiPs. These include, for example, TSMC's chip-on-wafer-on-substrate and Integrated Fan-Out (InFO), AMD's EPYC processors' technologies, and Intel's Ponte Vecchio, a high-performance compute GPU designed primarily for the exascale computing market, their EMIB (Embedded Multi-die Interconnect Bridge) and Foveros packaging technologies.

The examples are provided for tutorial purposes only and do not represent recommendations or endorsements from ETP. For more detailed information it is recommended that the companies are directly contacted

Chip-on-Wafer-on-Substrate

Chip-on-Wafer-on-Substrate (CoWoS) is an advanced packaging technology developed by Taiwan Semiconductor Manufacturing Company (TSMC). It's designed to address the challenges of integrating multiple chips into a single package to improve performance, reduce power consumption, and enable higher density. Here's a breakdown of CoWoS:

Key Features

1. **3D Integration**—CoWoS allows for the stacking of multiple chip layers (like logic chips and memory) in a single package, enhancing performance and reducing the distance that signals need to travel.
2. **Interposer Technology**—It uses a silicon interposer, a thin slice of silicon that acts as a bridge between the chips and the package. This interposer provides high-density interconnections and can help with thermal management and electrical performance.
3. **High Bandwidth**—By using the interposer, CoWoS provides high-bandwidth communication between chips. This is crucial for applications that require high data transfer rates, such as advanced computing and graphics.
4. **Fine-Pitch Technology**—The technology supports fine-pitch interconnects, allowing for a high density of connections and helping to reduce the overall package size.
5. **Enhanced Thermal Management**—The use of an interposer and the design of the CoWoS package can improve heat dissipation, which is essential for high-performance chips that generate significant amounts of heat.

Applications

- **High-Performance Computing**—CoWoS is often used in HPC applications where the integration of multiple high-performance chips can significantly boost computational power.
- **Graphics Processing Units**—For GPUs that require large amounts of memory and high-bandwidth interconnects, CoWoS provides an efficient solution.
- **Networking Equipment**—High-speed networking equipment benefits from CoWoS technology due to its ability to handle large volumes of data efficiently.
- **Advanced AI and Machine Learning**—CoWoS supports the high-bandwidth and processing requirements of AI and machine learning workloads.

Benefits

- **Performance Improvement**—By reducing signal travel distances and providing high-bandwidth interconnects, CoWoS can significantly boost overall performance.
- **Power Efficiency**—Efficient interconnects and thermal management can help reduce power consumption.
- **Compact Form Factor**—CoWoS technology can lead to more compact and integrated designs, saving space and potentially reducing costs.

Overall, CoWoS is a significant advancement in semiconductor packaging technology, enabling higher performance and more efficient integration of complex systems

Integrated Fan-out (InFO)

Integrated Fan-Out (InFO) is a semiconductor packaging technology developed by TSMC that provides a high-density, high-performance solution for advanced integrated circuits. This technology is designed to address the increasing demands for miniaturization, performance, and thermal management in modern electronic devices. Here's a detailed overview of InFO:

Overview of InFO Technology

1. **Concept**:
   - **Integrated Fan-Out (InFO)** is a type of fan-out wafer-level packaging (FOWLP) that integrates the IC die with the package substrate in a single process. It allows for a higher number of Input/Output (I/O) connections and better thermal and electrical performance compared to traditional packaging methods.
2. **Architecture**:
   - **Fan-Out Structure**—InFO uses a fan-out structure where the IC die is embedded in a mold compound, and the I/O connections are routed out to a larger area surrounding the die. This creates a fan-out area that extends beyond the edges of the die.
   - **Package Design**—The design often includes a redistribution layer (RDL) to route the electrical connections from the die to the external pads or balls on the package.
3. **Key Features**:
   - **High I/O Density**—By expanding the connections beyond the die's edges, InFO allows for a higher density of I/O pins, which is particularly beneficial for high-performance applications.
   - **Enhanced Performance**—The close proximity of the die to the package substrate and the use of RDL can improve signal integrity and reduce electrical interference.
   - **Compact Form Factor**—The fan-out design allows for a smaller overall package size while still providing a large number of I/O connections.

Benefits of InFO

1. **Improved Electrical Performance**:
   - **Reduced Signal Interference**—The direct routing of signals with short paths helps to minimize signal loss and interference.
   - **Higher Bandwidth**—The increased I/O density and improved signal routing contribute to higher bandwidth and faster data transfer rates.
2. **Enhanced Thermal Management**:
   - **Effective Heat Dissipation**—The design of InFO can include features like heat spreaders or thermal vias to improve heat dissipation and manage thermal performance effectively.
3. **Compact and Lightweight Design**:

- **Smaller Package Size**—InFO technology allows for a more compact package, which is advantageous for space-constrained applications in consumer electronics and other devices.

4. **Cost Efficiency**:
   - **Reduced Assembly Costs**—InFO's integration of die and package substrate in a single process can help reduce manufacturing and assembly costs compared to traditional packaging methods.
5. **Design Flexibility**:
   - **Integration of Different Technologies**—InFO supports the integration of various technologies, including logic, memory, and analog components, into a single package.

Applications of InFO

1. **Consumer Electronics**:
   - Used in smartphones, tablets, and wearables where high performance and compact size are critical.
2. **High-Performance Computing**:
   - Suitable for processors and memory modules that require high bandwidth and efficient thermal management.
3. **Graphics Processing Units**:
   - Applied in GPUs where high I/O density and performance are essential.
4. **Networking Equipment**:
   - Beneficial for networking devices that demand high-speed data transfer and compact design.
5. **Automotive Electronics**:
   - Used in automotive applications that require reliable and high-performance ICs in a small form factor.

Challenges and Considerations

1. **Manufacturing Complexity**:
   - The advanced manufacturing processes required for InFO can be complex and require precise control.
2. **Thermal Management**:
   - While InFO improves thermal performance, managing heat in high-density packages still requires careful design.
3. **Cost**:
   - The cost of advanced packaging technologies like InFO can be higher compared to traditional methods, although it may be offset by benefits in performance and miniaturization.

Overall, InFO represents a significant advancement in semiconductor packaging technology, offering improved performance, higher I/O density, and a more compact design for modern electronic applications.

AMD EPYC

AMD's EPYC processors are known for their high performance, energy efficiency, and scalability, making them ideal for data centers and enterprise applications. Below is an overview of the main features and chiplet architecture of AMD EPYC processors:

Main Features of AMD EPYC Processors:

1. **Zen Architecture**:
   - **Generations**—EPYC processors are based on AMD's Zen architecture, with iterations like Zen, Zen 2, Zen 3, and Zen 4. Each generation brings improvements in performance, power efficiency, and instruction sets.

- **Core Count**—EPYC processors are known for their high core counts, with models offering up to 96 cores per processor (in the EPYC 9004 series, for example).

2. **Infinity Fabric**:
   - **Interconnect Technology**—AMD's Infinity Fabric is a key technology that connects the various chiplets within the processor, enabling high-speed communication between cores, memory controllers, and I/O devices.
   - **Scalability**—This technology allows AMD to scale the number of cores and integrate various functionalities efficiently.
3. **Multi-Die Design (Chiplets)**:
   - **Chiplet Architecture**—EPYC processors use a chiplet-based design, where multiple smaller dies (chiplets) are connected. This approach improves yield, reduces costs, and allows for better thermal management.
   - **I/O Die**—The I/O die in the center handles memory and I/O, while the core chiplets (Compute Complex Dies or CCDs) contain the CPU cores.
4. **Memory Support**:
   - **DDR4/DDR5**—Depending on the generation, EPYC processors support either DDR4 or DDR5 memory, offering high memory bandwidth.
   - **Large Memory Capacity**—EPYC processors can support up to 4TB of memory per socket, which is crucial for memory-intensive applications.
5. **PCIe Support**:
   - **PCIe 4.0/5.0**—EPYC processors provide extensive PCIe lanes (up to 128 PCIe 4.0 lanes in EPYC 7003 series, for example), enabling high-speed connectivity for GPUs, storage, and networking devices.
6. **Security Features**:
   - **AMD Infinity Guard**—This suite includes features like Secure Encrypted Virtualization (SEV), Secure Memory Encryption (SME), and a hardware root of trust to protect data and workloads.
   - **Memory Encryption**—Real-time memory encryption helps protect against physical attacks on memory.
7. **Energy Efficiency**:
   - **Advanced Power Management**—EPYC processors incorporate power management features that optimize performance per watt, making them suitable for both high-performance and energy-efficient computing.
8. **High Performance**:
   - **Industry-Leading Performance**—EPYC processors often lead in industry benchmarks, offering exceptional performance in multi-threaded workloads like databases, virtualization, and high-performance computing.

Chiplet Analysis:

1. **Compute Chiplets (CCDs)**:
   - **Core Layout**—Each CCD typically contains up to 8 cores, with up to 12 CCDs in a single EPYC processor (as seen in the EPYC 9004 series).
   - **Cache Hierarchy**—Each CCD has its own L3 cache, which is shared among the cores, improving data access speeds for multi-threaded workloads.
2. **I/O Die**:
   - **Centralized I/O**—The I/O die handles memory controllers, PCIe lanes, and other I/O functions, centralizing these tasks to reduce latency and improve data throughput.
   - **Unified Memory Access**—This die also manages the memory access for all the compute chiplets, ensuring a balanced and efficient memory bandwidth across the processor.
3. **Scalability and Flexibility**:
   - **Multi-Socket Configurations**—EPYC processors can be used in multi-socket configurations, with up to 2 sockets per server, allowing for scalable solutions with up to 192 cores in a dual-socket system.

- **Modular Design**—The chiplet architecture provides AMD with the flexibility to mix and match different CCDs and I/O dies across different processor models, optimizing for various market segments.

Conclusion

AMD's EPYC processors leverage SiP by using a highly scalable chiplet architecture, advanced interconnect technology, and robust security features to deliver industry-leading performance and energy efficiency. These features make EPYC processors a popular choice for data centers, cloud computing, and enterprise applications.

Intel Ponte Vecchio

Intel's Ponte Vecchio is a high-performance compute GPU designed primarily for the exascale computing market. It's part of Intel's broader strategy in HPC (High-Performance Computing) and AI (Artificial Intelligence), particularly for applications in supercomputing. Ponte Vecchio is known for its innovative architecture, which incorporates a multi-tile or chiplet design, and advanced packaging technologies. Here's an overview of its main features and a detailed analysis of its chiplet architecture:

Main Features of Intel Ponte Vecchio:

1. **Xe-HPC Architecture**:
   - **Designed for HPC and AI**—Ponte Vecchio is based on Intel's Xe-HPC architecture, which is part of the larger Xe GPU architecture family. It's optimized for high-performance computing, AI workloads, and scientific simulations.
   - **Compute Power**—Ponte Vecchio is designed to deliver exceptional FP32 and FP64 performance, making it ideal for floating-point-heavy workloads such as those in scientific computing.
2. **Multi-Chiplet Design**:
   - **Tile-Based Architecture**—Ponte Vecchio uses a multi-chiplet (or multi-tile) design, where different functional blocks are implemented as separate chiplets. These chiplets are connected using advanced interconnect technologies.
   - **Heterogeneous Integration**—The chiplets include a mix of CPU, GPU, cache, and interconnect tiles, allowing for a highly customizable and scalable architecture.
3. **Advanced Packaging Technologies**:
   - **Foveros 3D Stacking**—Intel’s Foveros technology enables 3D stacking of chiplets, where one tile can be stacked on top of another, significantly increasing density and performance while reducing latency.
   - **EMIB (Embedded Multi-die Interconnect Bridge)**—EMIB is used for high-bandwidth, low-latency communication between chiplets, facilitating efficient data transfer across the GPU.
4. **High-Bandwidth Memory (HBM)**:
   - **HBM2e Integration**—Ponte Vecchio integrates High-Bandwidth Memory (HBM2e) directly on the package, providing enormous memory bandwidth crucial for HPC and AI workloads.
   - **Multiple HBM Stacks**—The design supports multiple HBM stacks, ensuring sufficient memory bandwidth to keep the GPU cores fed with data.
5. **Compute Tile Features**:
   - **Vector Engines**—Ponte Vecchio includes specialized vector engines optimized for dense matrix operations, essential for scientific computing and AI.
   - **Matrix Engines (XMX)**—These are designed for AI workloads, particularly for accelerating matrix multiplication tasks, which are common in deep learning.
6. **PCIe and CXL Support**:
   - **Connectivity**—Ponte Vecchio supports PCIe 5.0 and CXL (Compute Express Link) interfaces, offering high-speed connectivity with CPUs and other accelerators, reducing data transfer bottlenecks.

- **Scalability**—This connectivity allows Ponte Vecchio to be deployed in large-scale, distributed systems, making it suitable for exascale computing.

7. **Energy Efficiency**:
   - **Optimized Power Management**—Ponte Vecchio includes advanced power management features, optimizing performance per watt, which is critical in HPC environments where energy efficiency is paramount.
8. **AI and HPC Performance**:
   - **Tensor Operations**—Ponte Vecchio is optimized for tensor operations, providing high throughput for AI models, particularly in deep learning.
   - **Mixed-Precision Support**—It supports mixed-precision computations, allowing for a balance between performance and accuracy, depending on the workload requirements.

Chiplet Analysis:

1. **Compute Chiplets**:
   - **Core Architecture**—The compute chiplets contain the GPU cores, vector engines, and matrix engines. These are the workhorses of the GPU, performing the bulk of the computations.
   - **Scalability**—The chiplet approach allows Intel to scale the number of compute cores by adding more chiplets, enabling a customizable performance level depending on the target application.
2. **Cache and Memory Chiplets**:
   - **High-Bandwidth Cache**—Ponte Vecchio includes chiplets dedicated to cache, providing low-latency access to frequently used data and improving overall compute efficiency.
   - **HBM2e Chiplets**—The HBM2e memory is integrated as chiplets, providing high memory bandwidth directly on the package, reducing the need for external memory access and improving performance.
3. **Interconnect Chiplets**:
   - **EMIB and Foveros**—These technologies are used to connect the various chiplets, ensuring high bandwidth and low latency communication between the compute, cache, and memory chiplets.
   - **Unified Fabric**—The interconnect chiplets create a unified fabric that allows for seamless data transfer across the entire GPU, crucial for large-scale, parallel workloads.
4. **I/O Chiplets**:
   - **PCIe/CXL Interface**—The I/O chiplets handle the external interfaces, including PCIe 5.0 and CXL, facilitating high-speed communication with CPUs and other GPUs.
   - **Flexibility**—By separating the I/O functions into chiplets, Intel can customize the number and type of interfaces for different use cases.
5. **Packaging and Integration**:
   - **Modular Design**—The chiplet architecture allows Intel to mix and match different chiplets across various product lines, providing flexibility and efficiency in manufacturing.
   - **Thermal Management**—Advanced packaging helps in better heat dissipation, essential for maintaining performance in high-power HPC environments.

Conclusion

Intel's Ponte Vecchio represents a significant leap in GPU design, leveraging SiP with a multi-chiplet architecture and advanced packaging technologies like Foveros and EMIB. This design allows Intel to deliver a highly scalable and customizable GPU that meets the demanding requirements of exascale computing and AI. The integration of HBM2e, support for PCIe 5.0 and CXL, and specialized engines for AI and HPC workloads make Ponte Vecchio a critical component in Intel's HPC and AI strategy.

EMIB and Foveros

EMIB (Embedded Multi-die Interconnect Bridge) and Foveros are two advanced packaging technologies developed by Intel to enable high-performance, power-efficient, and scalable semiconductor designs. These

technologies are central to Intel's strategy of using chiplets (or multi-tile architectures) to build complex processors and accelerators like the Intel Ponte Vecchio. Here's a detailed analysis of each technology:

EMIB (Embedded Multi-die Interconnect Bridge):

Overview:

- **Purpose**—EMIB is a high-density, high-bandwidth interconnect technology that enables efficient communication between different chiplets on a single package without the need for large interposers or complex wiring schemes.
- **Implementation**—It involves embedding a small silicon bridge directly into the substrate of the processor package. This bridge connects the different chiplets, allowing them to communicate with each other with low latency and high bandwidth.

Key Features:

1. **High Bandwidth and Low Latency**:
    - **Efficient Data Transfer**—EMIB provides a direct, high-bandwidth pathway between chiplets, allowing for fast data transfer without the delays and energy consumption associated with longer interconnects.
    - **Reduced Signal Loss**—By minimizing the distance that signals need to travel between chiplets, EMIB reduces signal loss and noise, ensuring reliable communication.
2. **Scalability**:
    - **Flexible Chiplet Integration**—EMIB allows different chiplets (CPU, GPU, memory, I/O) to be integrated into a single package. This modular approach provides flexibility in combining various chiplets based on the needs of different products or applications.
    - **No Large Interposer Needed**—Unlike traditional 2.5D packaging, which requires a large silicon interposer to connect chiplets, EMIB uses small, localized bridges, which simplifies the manufacturing process and reduces costs.
3. **Power Efficiency**:
    - **Energy Savings**—EMIB's localized connections require less power to drive signals between chiplets, contributing to overall power efficiency in the package.
    - **Thermal Management**—The reduction in interconnect distance and the absence of a large interposer also aid in better thermal management, as heat can be more easily dissipated.

Applications:

- **HPC and AI Processors**—EMIB is particularly useful in high-performance computing and artificial intelligence applications, where multiple high-performance chiplets need to communicate efficiently.
- **Multi-Function Devices**—It enables the integration of diverse functionalities (e.g., compute, graphics, memory) into a single package, optimizing performance and space.

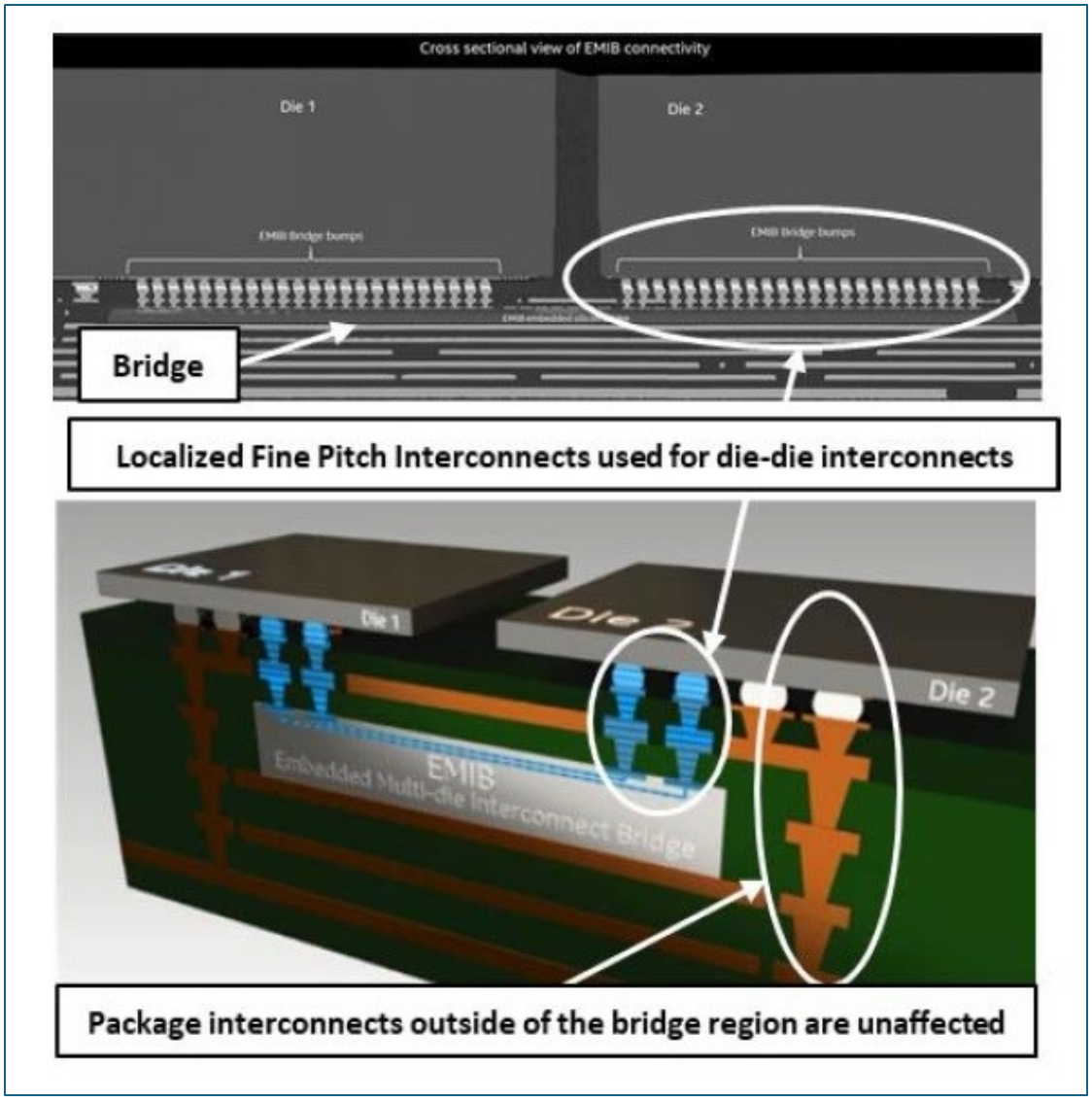


Figure A1. Example of EMIB interconnections

Source: Intel, Semi Wiki

Foveros (Greek for awesome)

Overview

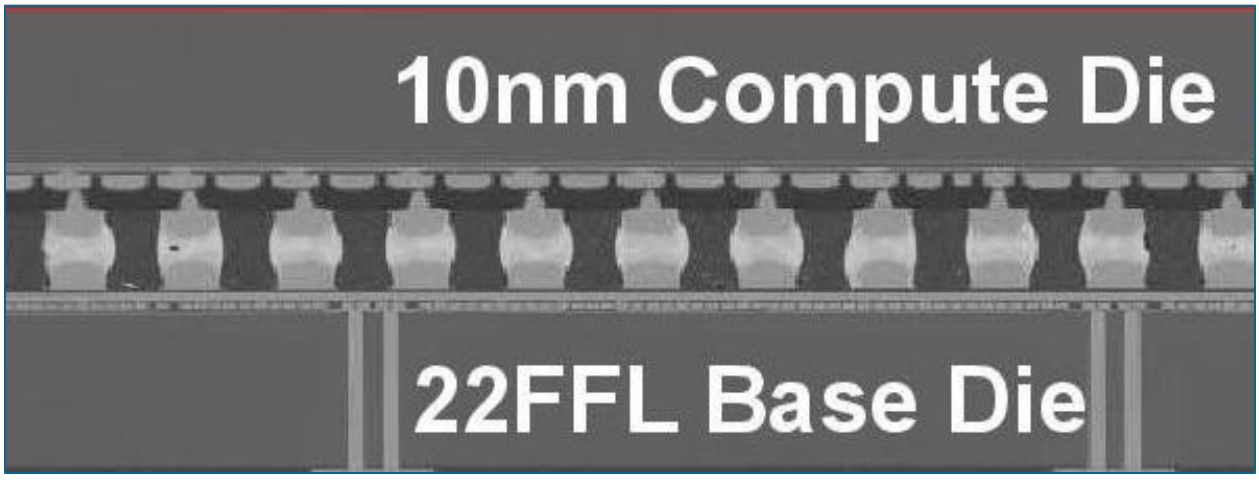
- **Purpose**—Foveros is Intel's 3D chip stacking technology that allows for the vertical integration of different chiplets or dies, effectively creating a three-dimensional (3D) package. This technology is critical for improving performance, increasing density, and enabling heterogeneous integration.
- **Implementation**—Foveros involves stacking different dies vertically, with through-silicon vias used to connect the stacked layers electrically. This 3D stacking approach allows for more compact designs with shorter interconnects between the layers.

Key Features:

1. **3D Stacking**:
    - **Vertical Integration**—Foveros allows multiple dies to be stacked on top of each other, which significantly reduces the footprint of the package. This vertical stacking also brings the different dies closer together, reducing interconnect length and improving performance.
    - **Heterogeneous Integration**—Foveros enables the combination of different types of dies (e.g., CPU, GPU, memory) in a single package, allowing for customized and highly optimized system designs.
2. **High-Density Packaging**:
    - **Increased Transistor Density**—By stacking dies, Foveros increases the transistor density of the package, enabling more powerful and feature-rich processors without increasing the overall size of the chip.
    - **Compact Designs**—The ability to stack dies vertically allows for more compact designs, which is particularly beneficial for mobile and space-constrained devices.
3. **Performance and Power Efficiency**:
    - **Reduced Interconnect Length**—The vertical stacking in Foveros significantly reduces the interconnect length between dies, resulting in lower latency and higher bandwidth communication.
    - **Power Efficiency**—The reduced interconnect length also contributes to lower power consumption, as less energy is needed to transmit signals between dies. Additionally, Foveros enables advanced power management techniques, allowing different layers to operate at different voltages or power states.
4. **Advanced Thermal Management**:
    - **Heat Dissipation**—Although 3D stacking can lead to increased heat density, Intel has implemented advanced thermal management techniques within Foveros to efficiently dissipate heat. This includes using thermal interface materials and optimizing the placement of hot and cool components within the stack.

Applications:

- **High-Performance Processors**—Foveros is ideal for processors that require high performance in a small form factor, such as those used in high-end laptops, AI accelerators, and data center applications.
- **Heterogeneous Systems**—It supports the development of heterogeneous computing systems where different types of compute units (e.g., general-purpose cores, AI engines, GPUs) are integrated into a single package for optimized performance.



Source: Intel

Figure A2. Foveros interconnections allowing full 3D die stacking

Source: Intel

Comparative Analysis:

1. **Interconnect vs. Stacking**:
    - **EMIB** focuses on horizontal integration, connecting chiplets side by side within a package. It is ideal for designs where chiplets need to communicate across a 2D plane.
    - **Foveros**, on the other hand, focuses on vertical integration, stacking different dies on top of each other. This is beneficial for increasing density and reducing the footprint of the package.
2. **Scalability and Flexibility**:
    - **EMIB** provides excellent scalability in terms of adding more chiplets horizontally, making it easier to integrate a variety of functionalities (e.g., compute, memory, I/O) in a single package.
    - **Foveros** offers flexibility in stacking different types of dies vertically, enabling more compact and power-efficient designs, particularly in products where space is at a premium.
3. **Performance and Power**:
    - **EMIB** excels in reducing latency and power consumption for horizontal communication between chiplets, making it ideal for high-performance computing environments.
    - **Foveros** reduces interconnect lengths vertically, which can further enhance performance and power efficiency, especially in applications requiring high transistor density and compact form factors.

Conclusion

Both EMIB and Foveros are pivotal in Intel's SiP chiplet strategy, enabling the creation of powerful, flexible, and scalable processors. EMIB provides a robust solution for horizontal chiplet integration, offering high bandwidth and low latency communication, while Foveros allows for vertical stacking, increasing transistor density and enabling heterogeneous integration. Together, these technologies empower Intel to design and manufacture cutting-edge processors for a wide range of applications, from HPC and AI to mobile computing and data centers.