THE HETEROGENEOUS INTEGRATION ROADMAP: ENABLING TECHNOLOGY FOR SYSTEMS OF THE FUTURE

Paul Wesling, IEEE Life Fellow
Hewlett-Packard Co (retired); HI Roadmap Editor
p.wesling@ieee.org

ABSTRACT
The new Heterogeneous Integration Roadmap (HIR) provides a long-term vision for the electronics industry, identifying difficult future challenges and potential solutions. Under the sponsorship of SEMI, ASME, and three IEEE Societies, the roadmap offers professionals, industry, academia, and research institutes a comprehensive view of the landscape and strategic technology requirements for the electronics industry's evolution over the next 15 years, and provides a 25-year vision for the heterogeneous integration of emerging devices and emerging materials with longer research and development timelines. The purpose is to stimulate precompetitive collaboration and thereby accelerate the pace of progress.

The International Technology Roadmap for Semiconductors (ITRS) set the cadence for the Moore's Law scaling that has been the norm for the semiconductor industry. However, because of scaling, cost and power-dissipation issues, as well as the laws of physics, the final ITRS was issued in 2015. The HIR pulls together many strands of that earlier Roadmap, to focus on microelectronics design, materials and packaging issues. The current version covers 2.5D, 3D, and wafer-level packaging, integrated photonics, MEMS and sensors, and system-in-package (SiP); support areas such as test, thermal, simulation, co-design, and interconnects; as well as application areas such as high-performance computing, 5G, medical, aerospace, automotive, and mobile – detailing both near-term and longer-term metrics and goals. It identifies difficult future challenges and proposes potential solutions. Comprising the output of 22 Technical Working Groups with worldwide participation, it will be substantially updated every two years.

Version 1.0 is available freely for download, as well as in the form of a printed softbound book. Details for accessing this new Roadmap are presented. An invitation is made for involvement in version 2.0, now under preparation.

INTRODUCTION
Electronics are deeply embedded into the fabric of our society, changing the way we live, work and play while bringing new efficiencies to our global lifestyles, industries, and businesses. We are in the era of the digital economy and ubiquitous connectivity, and the market forces driving data and systems growth include:

- Migration of data, logic, and applications to the cloud
- Consumerization of IT with the rise of social media
- Evolution in our mobile devices
- 5G communications together with Internet of Things (IoT) to Internet of Everything
- Artificial Intelligence (AI) with Virtual Reality (VR) and Augmented Reality (AR)
- Autonomous vehicles

Heterogeneous integration refers to the integration of separately manufactured components into a higher-level assembly (System in Package – SiP) that, in the aggregate, provides enhanced functionality and improved operating characteristics. In this definition, components should be taken to mean any unit, whether individual die, MEMS device, passive component or assembled package or sub-system, that are integrated into a single package. The operating characteristics should also be taken in its broadest meaning including characteristics such as system-level performance and cost of ownership.[1]

Since the 1970's, Moore's Law has predicted the continued increase in semiconductor areal density with a corresponding reduction in voltage, increase in switching speed, and decrease in costs. However, Moore's Law economics are coming to an end and some key performance metrics at advanced nodes are plateauing.[2]

Heterogeneous integration follows naturally from the conceptual vision of building large, complex systems out of smaller functions separately manufactured and packaged. Heterogeneity and the associated integration are far-reaching and can relate to materials, component types, circuit types, semiconductor nodes, interconnect methods, and source or origin. There are many examples of heterogeneous integration today, where a full subsystem or system cannot yet be fully integrated into a single semiconductor device (system-on-chip: SOC). Examples include the Intel Agilex and the AMD EPYC applications (Figures 1, 2). This direction is also evident in DARPA's CHIPS program (Figure 3).

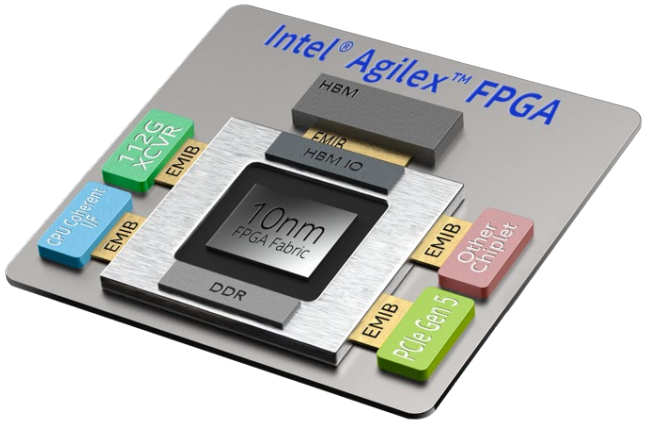

Figure 1. Intel Agilex FPGA Chiplet application. Source: Intel

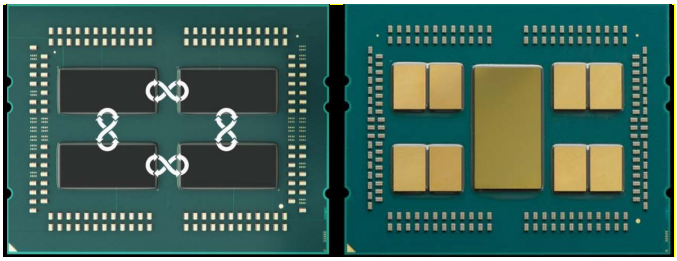

Figure 2. AMD EPYC Server Processors. Source: AMD

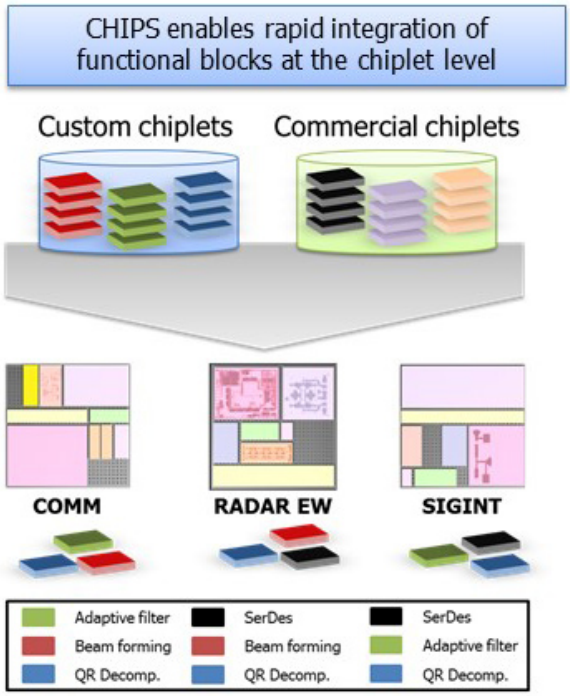

Figure 3. DARPA CHIPS Program. Source: DARPA

A key driving market for HI is the smartphone because of its requirement for a thin profile, low power consumption, and a drive for lower costs. This heterogeneous modularity allows for upgrading various internal devices (such as incorporating a main processor from the next node) while continuing to integrate other devices from an earlier node (such as memory, MEMS and RF). Examples from Apple,

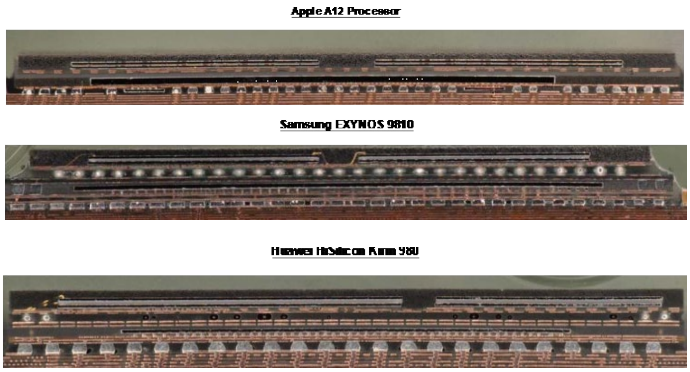

Figure 4. Three examples of Package on Package [PoP] in Smart Phone Teardowns.

Source: Prismark Partners & Binghamton University

Samsung and Huawei are shown in Figure 4. Improved versions of this figure and other examples are included in Chapters 1 and 8 of the Roadmap.

HIGH-PERFORMANCE COMPUTING AND DATA CENTERS

Semiconductor devices targeting the high-performance computing (HPC) and data center markets have always represented the best state-of-the-art in devices and process technologies. The needs in these market segments have generally demanded the highest processing rates, highest communication rates (low latencies and high bandwidth, often both simultaneously) and highest capacities, in addition to requirements for packaging that address the interconnection requirements and higher-power dissipation. This is a trend that is likely to continue for a wide variety of applications for HPC systems and data centers over the coming years.

A chapter in the new Roadmap lays out the need for heterogeneous system integration to realize systems-in-a-package (SiPs), and identifies potential solutions and short-term, medium term and longer-term challenges that are encountered in realizing these SiPs. Yield issues, device limitations and escalating design costs make it uneconomical to maintain the die size with technology scaling. Consequently, even though circuit elements are still scaling, performance scaling is no longer happening following Moore's Law. Heterogeneous integration offers a solution for performance scaling. Instead of fabricating a single large multicore CPU die, smaller dies can be tiled within a package on an interposer with very short connections between the dies to realize the same performance offered by a single large die. The smaller dies have higher yields and as long as integration cost is reasonable, the overall SiP solution scales in performance following Moore's Law.

Multicore CPUs and GPUs impose a severe demand on the memory system in terms of both latency and particularly bandwidth. Without the availability of low-latency, high-

bandwidth connections to memory, the performance potential of these processing engines remain unexploited. Placing memory off-chip, away from the processing elements and limited by package IO parameters, precludes the realization of low latency, high bandwidth connections. Placing the memory within the package forms an additional layer within the memory hierarchy and is supplemented with additional memory outside the heterogeneously integrated package.

High bandwidth memory (HBM) takes the approach of stacking memory dies and using wide-channel interconnections. The HBM technology thus keeps the real estate needs of the DRAM fixed and grows capacity by increasing the stacking count. The SRAM or DRAM memory dies can be placed adjacent to the processor on an interposer (Figure 5) or stacked onto the processor. A JEDEC standard now defines five successive generations of HBMs.[3] This future direction and a number of packaging options are covered in detail in the new Roadmap.

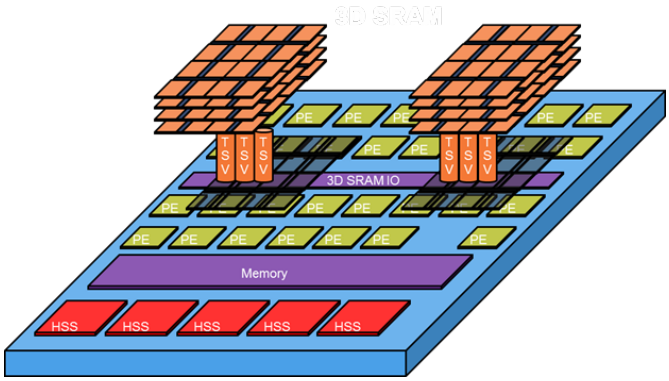

Figure 5. Integration of stacked SRAM

In the short term (2 to 5 years), potential solutions include ultra-large silicon interposers, embedded-die bridges in both silicon and glass, TSV scaling for fine-pitch wiring between die levels, face-to-face stacking, and hybrid bonding. Photonics will soon play a significant role in both enabling the use of high-end SiPs for HPC and data centers into the rest of the system as package IO solution in the near term, and ultimately to facilitate tighter integration in high-end SiPs in the longer term. The limitations of copper as an interconnection material will be addressed for SiPs for these applications, where low interconnection latency is important. The RC delays of copper can no longer be ignored as nodes shrink below 3nm. Solutions to stretch the use of copper connections include annealing, use of shielding materials, and other approaches.

AUTOMOTIVE AND VEHICULAR TECHNOLOGY

The next few years will bring key disruptive trends in vehicular electronics. The electronics content in a car has been continuously growing, and is expected to increase in 2030 to about 50% of the total cost of an automobile. The increased emphasis on autonomous driving as well as electrification of vehicles has resulted in enormous changes for semiconductors and batteries used and their packaging through heterogeneous integration for next-generation automobiles and trucks. Trends in automotive electronics include autonomous driving, in-car smartphone-like infotainment, privacy and security, low-latency communication (5G), and all-electric cars, requiring enhancements in current semiconductor and packaging technologies as well as an entirely different set of technologies.

This Roadmap chapter introduces highly complex packaging for processors used in autonomous driving, integration of advanced communications, and the associated challenges with ensuring higher levels of reliability in all components based on new use cases for automobiles and general transportation going forward. Numerous advances are expected in sensor technology with advancements of sensing techniques. Integration of power systems will continue as cars continue to electrify. Artificial Intelligence (AI) will be central to both the functionality and safety of vehicles, as well as in techniques used for advancing reliability of the electronic components. The state of the art is covered and focuses on 5-, 10- and 15-year roadmap needs, challenges, and potential solutions in the three main drivers for automotive electronics.

Sensing technologies are key elements to enable migration to autonomous vehicles. The broad categories of sensors in autonomous vehicles include cameras, RADAR, LiDAR, and ultrasonics. Combinations of various sensors are needed for the fully autonomous vehicle. Cameras provide vision, but not completely. RADAR, LiDAR and ultrasonics are also essential for enabling autonomous driving. Autonomous driving demands highly robust sensing technologies all around the car for real-time detection, surround-view, and collision avoidance. By integrating the information obtained from all these sensor systems and processing the data in real time, through machine learning and artificial intelligence, fully self-driven autonomous vehicles can be realized.

5G COMMUNICATIONS

Fifth-generation mobile communications (5G) represents new opportunities for IC packaging technology that are significantly different from previous generations of cellular-based mobile wireless technologies (Figure 6). Identified in this Roadmap are challenges at 5-, 10-, and 15-year horizons along with guidance on how to meet those challenges.

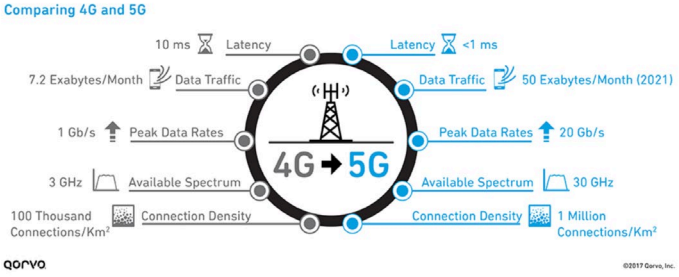

Figure 6. Comparing 4G and 5G[4]

The primary focus is for millimeter-wave 5G bands, which provide the greatest challenge to packaging of RF circuits and the necessary antennas. Addressed are both base-station

equipment and user equipment (UE – eg, cellphones and autonomous vehicles). To achieve millimeter-wave high-bandwidth connectivity, 5G systems will utilize beamforming technologies. Instead of wide-angle radiation patterns, narrow-focused beams will be used to track the UE devices as they move around within the operating radius of the cell site. This kind of phased-array technology has been used previously in radar systems for tracking targets. Now with 5G, phased-array beamformers will be used at higher powers at the base-station and at lower power levels at the UE. There are three variants of beamforming: analog, digital and hybrid. Analog beamforming uses analog techniques to alter the amplitude and phase of the signal leaving each radiating element. Analog beamforming is simpler and more power-efficient as compared to digital beamforming. Digital beamforming is able to support multiple simultaneous beams but requires more power consumption due to analog-to-digital conversion needed at each element. The hybrid architecture is well suited for large arrays and has perhaps the optimum architecture for massive MIMO (256 elements or more).

Challenges include the small element-to-element spacing needed for the phased-array; choosing the right semiconductor technology based on the needed output power; and the integration level, multiple-band coverage, and packaging required. An implementation of 5G in a cellphone shows the three different RF/antenna implementations currently needed to provide adequate 5G reception for a single mobile device. Details are shown in the Roadmap.

ADDITIONAL ROADMAP COVERAGE

While several technology trends have been discussed above, the Roadmap also covers application areas such as medical and health, Internet of Things (IoT), aerospace and defense, and mobile. Chapters on supporting technologies cover thermal, 2D and 3D architectures, co-design, modeling and simulation, SiP and module design, wafer-level packaging, single-chip and multi-chip packaging, integrated photonics, MEMS and sensors, and integrated power electronics. Other areas are covered, such as reliability, test, supply chain issues, security, and emerging research devices and materials. Graphs and tables cover pre-competitive technology requirements at three-year intervals through 2032.

SUMMARY

Electronics packaging is fundamental to innovation in many industry applications and system products, across applications such as cloud computing, IoT, medical and health, automotive, aerospace, consumer, and home. The HI roadmap offers professionals, industry, academia, and research institutes a comprehensive view of the landscape and strategic forecast of technology requirements for the electronics industry's evolution over the next 15 years. The goal is to stimulate our total ecosystem towards precompetitive collaboration and thereby accelerate the pace of progress.

Work is now underway on version 2.0 of the Roadmap. Technologists are invited to become involved in its preparation; visit eps.ieee.org/hir for a summary of the various Technical Working Groups and for contact information.

The full Roadmap is available as a single PDF download, and also as a ZIP file to place in a directory on a desktop or laptop computer, with full-text search enabled; download at www.pwesling.com/hir. A reference link is given there for purchase of the softbound book. Individual chapters (with updates as they become available) may be downloaded from the IEEE Electronics Packaging Society website, at eps.ieee.org/hir.

REFERENCES

1. International Technology Roadmap for Semiconductors (ITRS), Heterogeneous Integration Chapter, http://www.itrs2.net/
2. Moore's Law Saturation, *The Economist*, March 12, 2016
3. High Bandwidth Memory, Wikipedia pages at: https://en.wikipedia.org/wiki/High_Bandwidth_Memory
4. "A Brief History of Mobile Communications," found at http://www.winlab.rutgers.edu/~narayan/Course/Wireless_Revolution/vts%20article.pdf