Advanced Packaging for Heterogeneous Integration

Ravi Mahajan
Intel Fellow
Assembly and Test Technology Development
Chandler, AZ

Legal Notices

This presentation contains information provided by Intel Corporation ("Intel"), and may refer to Intel's plans and expectations for the future, which are provided for discussion purposes only and are subject to change without notice. Forward-looking statements involve a number of risks and uncertainties: Refer to Intel's SEC filings for authoritative discussion of Intel's results and plans. This presentation imposes no obligation upon Intel to make any purchase, and Intel accepts no duty to update this presentation based on more current information. Intel is not liable for any damages, direct or indirect, consequential or otherwise, that may arise, directly or indirectly, from the use or misuse of the information in this presentation.

Copyright © 2019 Intel Corporation. Intel and the Intel logo, are trademarks of Intel Corporation in the U.S. and/or other countries. Other names and brands may be claimed as the property of others.

Outline

- Brief History of Intel Innovations
- Why is Advanced Packaging so Important?
- Intel's Current Advanced Packaging Portfolio
- A Look into the Future

Intel has been on the forefront of many “Quiet” Revolutions in Packaging Technology

- 1990s
  - Packaging revolutionized by launch of the first flip-chip organic substrate
  - First Organic Pin Grid Array launched – A New Package Industry is born
  - > 100W Cooling Solutions developed; Laptop Heat pipes become mainstream
- 2000s
  - Intel introduces Copper Pillar Bumps on die
  - Intel leads the industry by deploying fully lead & halogen free packages
  - Embedded Array Capacitors implemented for HPC – another first by Intel
  - High Pin Count Land Grid Sockets Become Mainstream
  - Drive to greater adoption of Ball Grid Arrays enables “Small and Thin”
  - >200W Cooling
- 2010s
  - Thermo-Compression Bonding Tools enable Fine Pitch Flip-Chip
  - EMIB – Breakthrough 2D Multi-Chip Packaging Introduced
  - Foveros opens up the 3rd Dimension

FUTURE FAB International Issue 17 - PROCESS GASES, CHEMICALS AND MATERIALS

The Mechanical Side Of Ultra-Low k: Can It Take the Strain?

Special ULK Focus

The need for increased performance and reduced power in large logic chips has driven reduction of the dielectric constant for the interlayer dielectric (ILD) layers with each succeeding silicon generation. The primary approach to reduce dielectric constant has been the introduction of increasing amounts of porosity in the ILD material. A consequence of this has been the increasing fragility of the ILD (typically referred to in the literature as low k or ultra-low k (ULK), resulting in increased reliability concerns driven first by the CTE mismatch between die and package and the transfer of stresses from the package to the die through the die-package connection, and second by the impact of packaging processes on the mechanically weak ILD. This paper provides an overview of the impact of the package on the ILD, as well as mechanical requirements for low-k ILD materials to address reliability concerns.

Biju Chandran
Ravi Mahajan
Mark Bohr
Chia-Hong Jan
Quat T. Vu

The microelectronics industry has been relentlessly driven by Moore's Law,[1] a corollary of which is increasing chip performance. Microprocessor performance itself depends on many factors besides silicon technology, key among these being the architecture. A component of silicon technology and architectural performance is the interconnect performance, usually measured by a simple metric, i.e., the RC constant, R and C being the resistance and capacitance of the interconnect respectively. The dielectric constant (k) of the ILD layers comes in directly in the factor C of this metric. With each succeeding generation of silicon technology starting with 130 nm node, there has been an effort to reduce the RC constant. The transition from aluminum interconnects to copper-based interconnects provided a significant opportunity to reduce R. Since the Al to Cu transition, the focus has been to reduce C. It is expected that starting with 45 nm node, dielectric constant reduction will result from the introduction of porosity into the ILD. Reduction in dielectric constant of the ILD has been accompanied by a reduction in its adhesive and cohesive strength. ILD layers have become increasingly fragile with each silicon generation, leading to increasing reliability concerns. Package designs and die-package interconnect increasingly impact ILD reliability. In the past few years, we have seen an increased level of interest in the study of die package interactions as they impact reliability and performance.[2,3,4] The focus of this paper is to discuss the ILD mechanical reliability challenges at 45 nm and beyond due to die-package interactions.

ILD Materials and Mechanical Strength

ILD Material Directions

Figure 1 shows the ILD stack-up in a typical application and is used to set the context for the discussion that follows. Advanced microprocessors depend upon transistor scaling to meet density and performance targets. The accompanying interconnect scaling causes the interconnect propagation delays to be a significant portion of the clock cycle time, impacting the chip performance.[5] The use of copper instead of conventional aluminum interconnects is the first step toward reducing this interconnect delay by reducing the resistance of the interconnects. Further reduction in interconnect delay is achieved by reducing the capacitance in the ILD through the introduction of materials with low dielectric constants (Figure 2).

An unintended consequence of driving for lower k is that ILD materials become increasingly fragile, as seen in Figure 3. In a

Figure 1. Cross-sections showing the die back-end structures in an SRAM.

FUTURE FAB International

Thermo-compression Bonding for Fine-pitch Copper-pillar Flip-chip Interconnect – Tool Features as Enablers of Unique Technology

Amram Eitan; Kin-Yik Hung
Intel Corporation; ASM Pacific Technology Ltd.
5000 W. Chandler Blvd., AZ, USA; 16-22 Kung Yip Street, Kwai Chung, Hong Kong
Amram.eitan@intel.com; kyhung@asmpt.com

intel

2D AND 3D PACKAGING DRIVE NEW DESIGN FLEXIBILITY

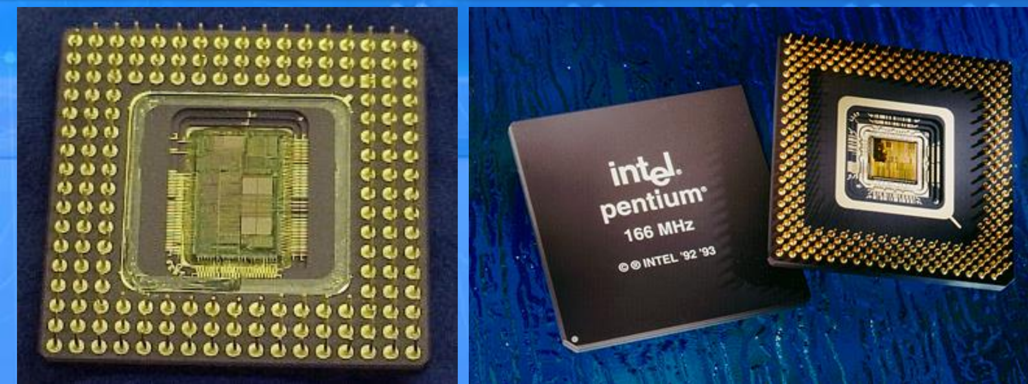
The combination of advanced 2D and 3D packaging technologies allows Intel to flexibly combine smaller chiplets of IP to meet the demands of a huge range of applications, power envelopes, and form factors. Intel® embedded multi-die interconnect bridge (EMIB) and Foveros are advanced 2D and 3D packaging technologies, delivering high performance at low cost.

MONOLITHIC	2D INTEGRATION	3D INTEGRATION
Integrate functions on a single die for high performance on a single silicon technology	Combine IPs built with separate processes into a single package with Intel EMIB, helping improve yield, cost, time-to-market, and total capability	All the benefits of 2D integration plus a new level of density thanks to Foveros, allowing for a radical re-architecture of systems-on-chips

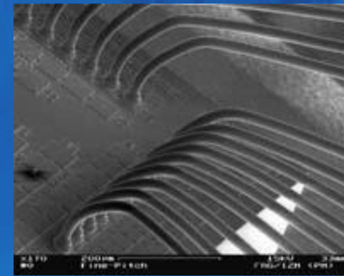
GRAPHICS/IMAGING
COMMS
CPU CORES
MEMORY

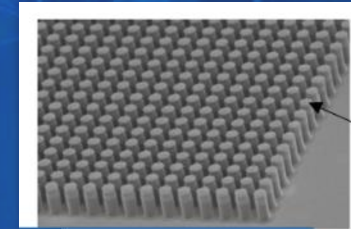
I will focus Mainly on Major Changes in Packaging Interconnects

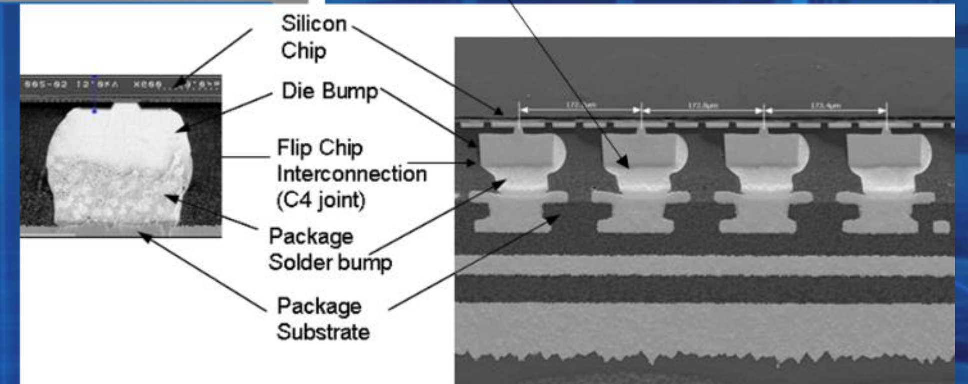
In the 1990s : Transition to Organic Flip-Chip……



+





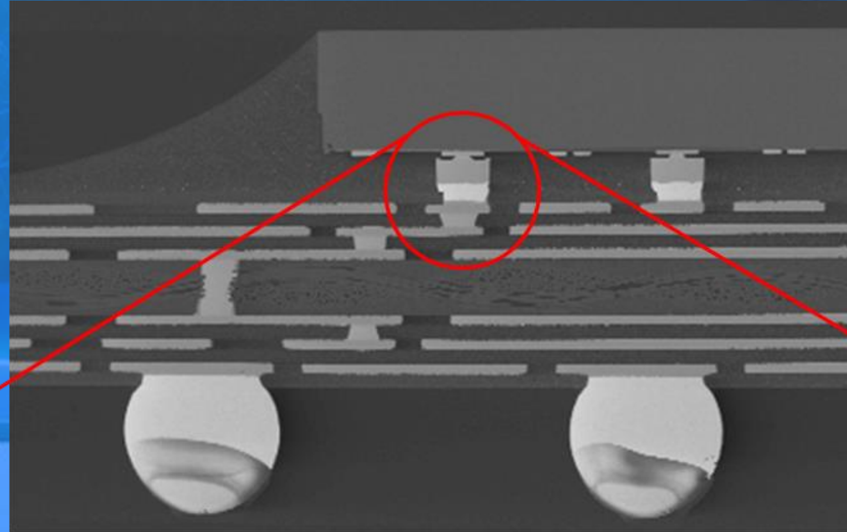


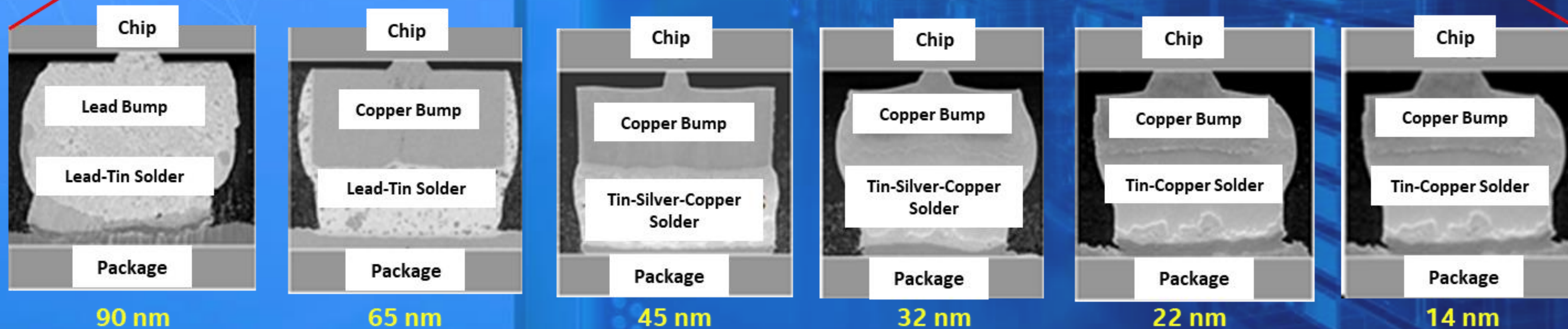
Ceramic Packaging → Organic Packaging

Wire-Bond → Organic Flip-Chip

Transition to Low Cost, Copper Based, Organic Area Array Packaging Enhanced CPU Performance with Improved Signaling and Power Delivery

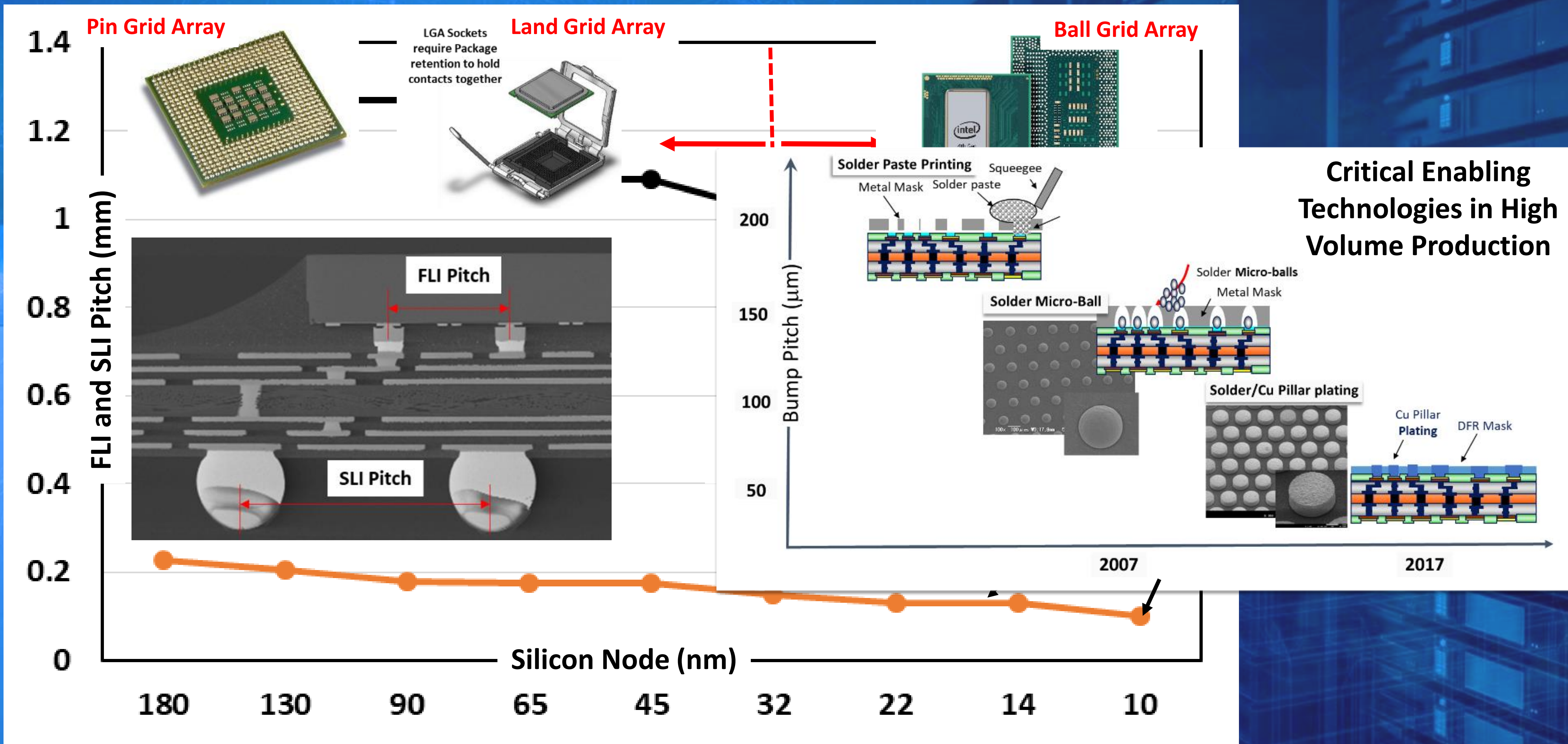
In the 2000s.....The Lead-Free Transition





Intel was the First to Completely Eliminate Lead from the Package

All Along Interconnects Continued to Scale



Packaging is for Ideal for Heterogeneous Integration*

Moore's Law,
40 years and Counting
Future Directions of Silicon and Packaging

Bill Holt
General Manager
Technology and Manufacturing Group
Intel Corporation

InterPACK '05
2005 Heat Transfer Conference

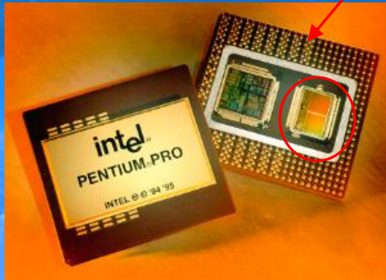
Key Messages

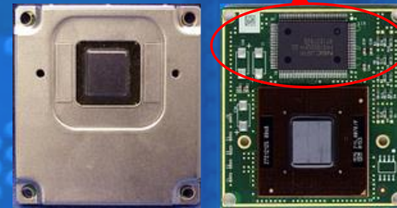
- Systems/platform focus drives increased requirements for silicon and packaging
- Making the right choice between on-chip and in/on package integration is critical to cost effective solutions
- Moore's Law is the engine for continued growth
- Silicon is ideal for homogenous integration
- Packaging is ideal for heterogeneous integration

*Heterogeneous Integration is the integration of separately manufactured & tested components into a higher level assembly (SiP aka MCP) that, in the aggregate, provides enhanced functionality and improved operating characteristics

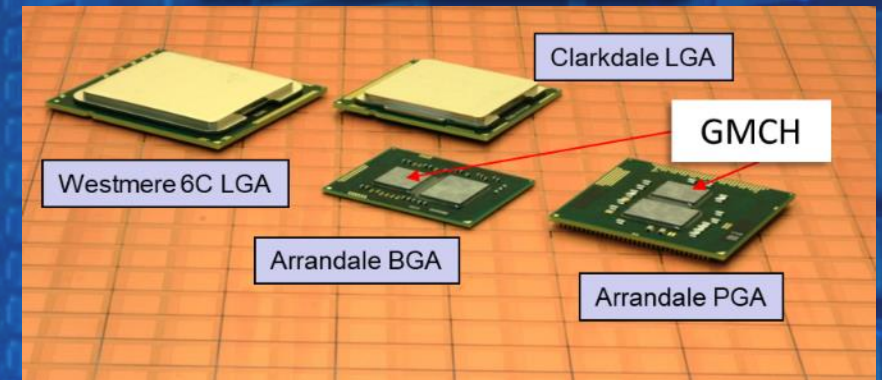
Intel* has a long history of using MCPs for (TTM) & Performance….

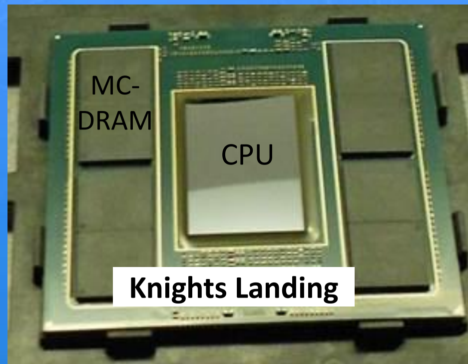
SRAM Integration

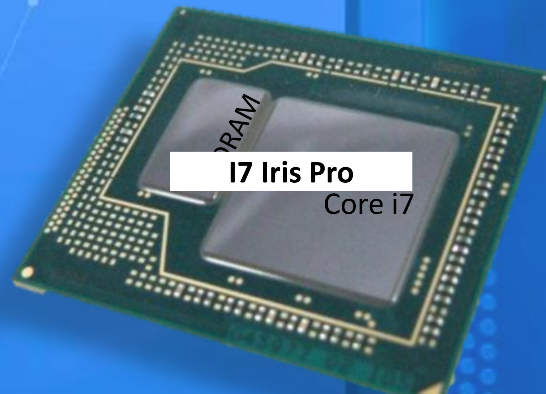


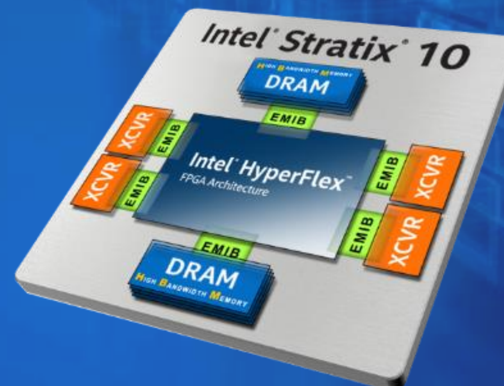


(G)MCH Integration









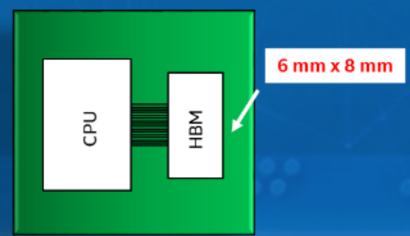
DRAM Integration

* Other companies (e.g. IBM, AMD, NVidia) have also similarly employed MCPs

Why is Advanced Packaging So Important?

The Package as a Compact Integration Platform

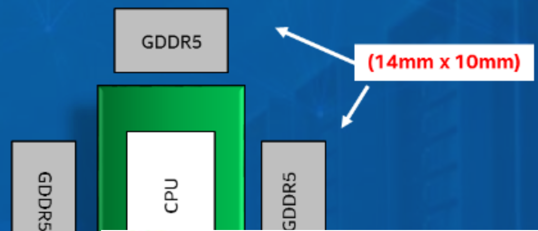
2016-2017



HBM

- Total Capacity 4GB (1GB each)
- Data rate – 1 - 2Gb/s
- Total BW – (128-256) GB/s
- IO Power Efficiency (Energy/bit) – 1X

1. http://www.memcon.com/pdfs/proceedings2015/MKT105_SKhynix.pdf
2. https://www.micron.com/~/media/documents/products/technical-note/dram/tned02_gddr5x.pdf



- Total Capacity
- Data rate – 12
- Total BW – 19
- IO Power Effic

2017-2019

Memory System Comparison: 256GB/s GDDR6 vs. HBM2

GDDR6 – 14mm x 12mm

GDDR6 Memory System: Four 16Gbps x32 GDDR6 DRAMs

HBM2 Memory System: Single 2Gbps HBM2 Device

Narrow and Fast

Wide and Slow

	GDDR6 Memory System	HBM2 Memory System	
Total Bandwidth	256 GB/s	256 GB/s	
Per-pin data rate	16 Gbps	2 Gbps	
Relative Controller PHY Area[1]	1.5-1.75	1.0	Area advantage for HBM2
Relative Controller PHY Power[1]	3.5-4.5	1.0	Power advantage for HBM2
Interposer	None	Added cost[2]	Cost and complexity advantage for GDDR6
Memory	Similar to GDDR5, DDR4	Stacked, adds cost[2]	Cost advantage for GDDR6

[1] Source: Rambus Inc.
[2] Source: The Cost of HBM2 vs. GDDR5 & Why AMD Had to Use It, https://www.gamersnexus.net/guides/3032-vega-56-cost-of-hbm2-and-necessity-to-use-it

https://www.rambus.com/blogs/key-design-points-to-consider-gddr6-and-hbm2-drams/

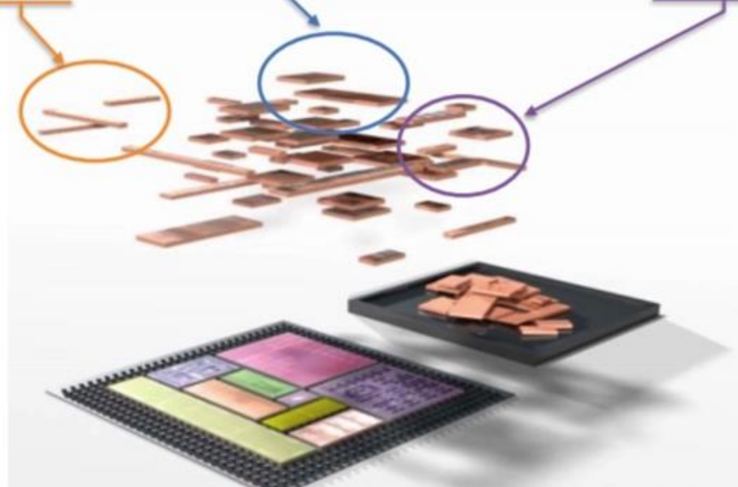
On-Package Integration is More Compact, Lower Power & Higher BW

Recent Interest in Advanced Packaging is driven by….

10000

Need for High Bandwidth*

2018 2020 2022



Need to Integrate IP on different Nodes & Fabs

CHIPS modularity targets the enabling of a wide range of custom solutions

Yield Resiliency

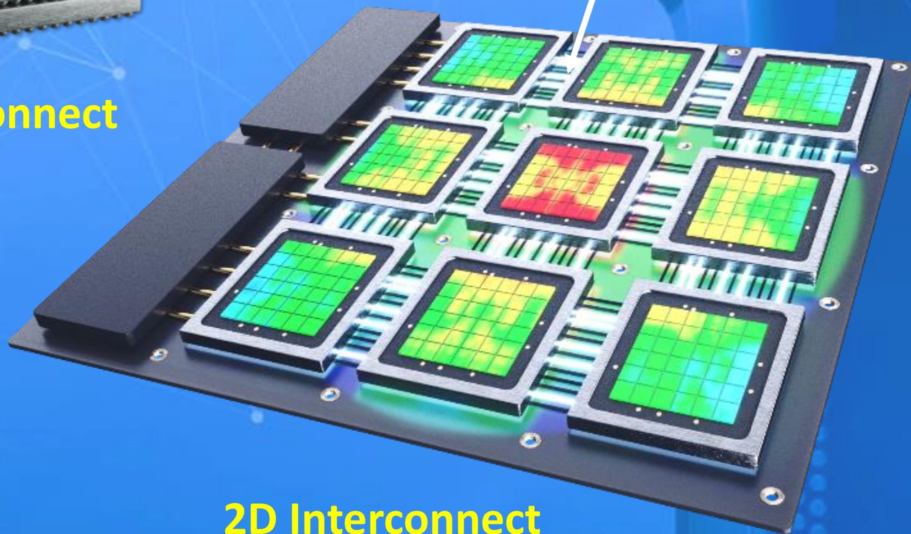
$$Die\ Yield \approx exp\left[-\left(\mathbf{Die\ Area}\ x\ \frac{Process\ Defects}{Area}\right)\right]$$

* Intel trends for on-package memory BW

Interconnects in Advanced Packaging….

High Bandwidth, Low Power, Parallel Links drive Need for High Density Die-Die Interconnects

3D Interconnect



2D Interconnect

The Serial vs. Parallel Question

Source: Presentation by Andreas Olofsson @ DARPA 2.5D/3D Workshop, Dec 6, 2018

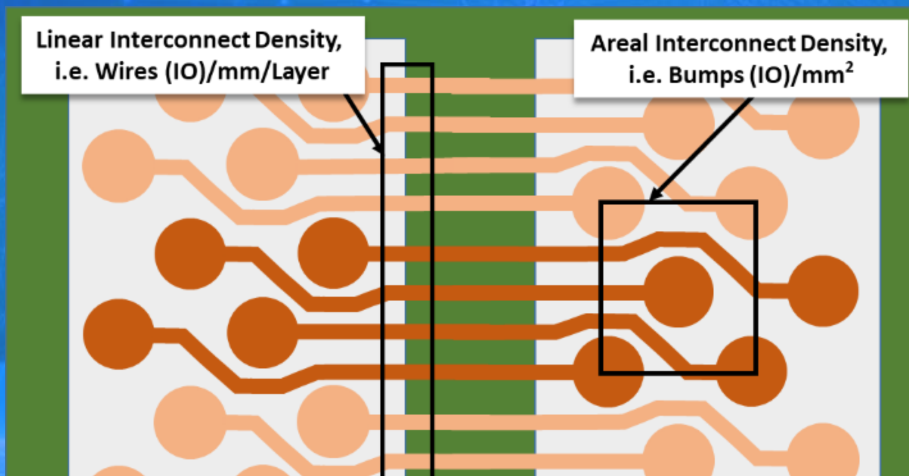
	CHIPS Parallel		Serial
IP Complexity	Flip-flop + tristate		SERDES
IP Cost	Low		High (open source?)
Throughput	1Tbps/mm		1Tbps/mm?
Latency	<5ns	Physics	High
Energy Efficiency (<1000um)	0.1pJ/bit		1-10pJ/bit
Throughput per pin	2Gbps		30Gbps
Packaging Complexity	High	Economics?	Low
Packaging Cost	High		Low

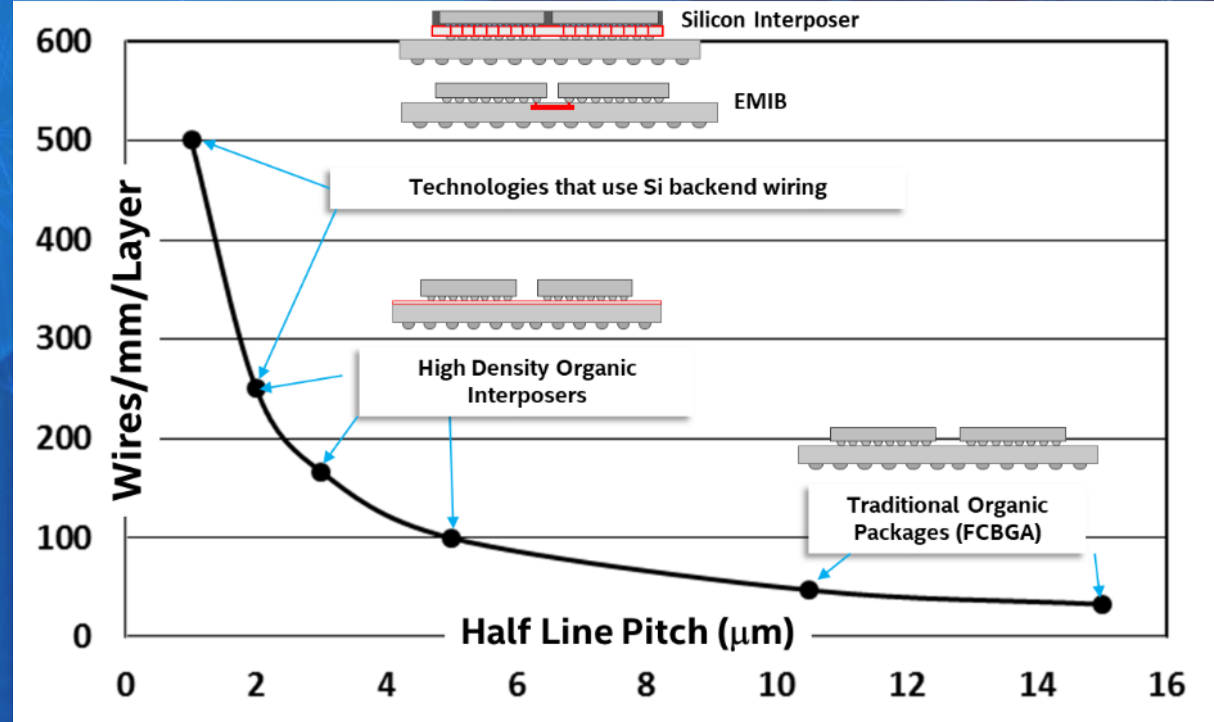
Parallel will continue to be the chosen path for DARPA until someone makes a valid case for a different option.

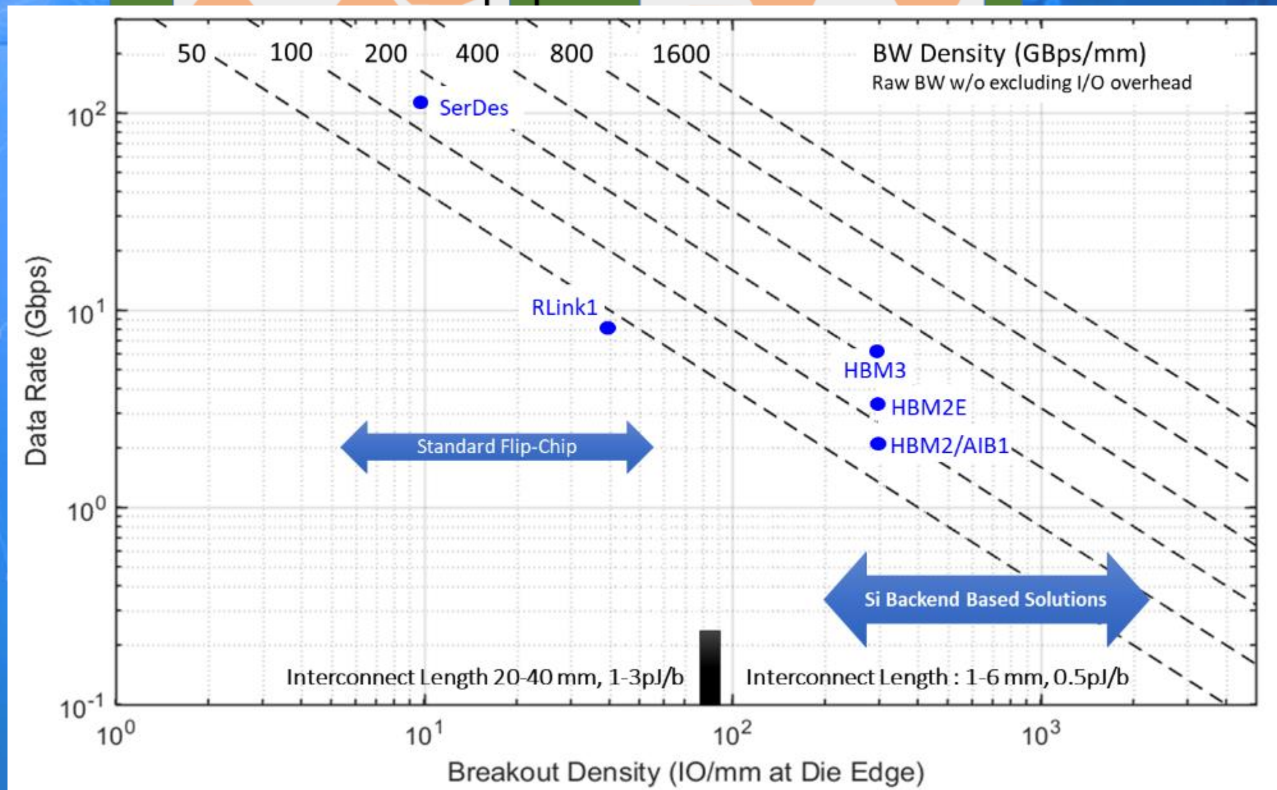
Distribution Statement "A" (Approved for Public Release, Distribution Unlimited)

11

2D MCP Landscape

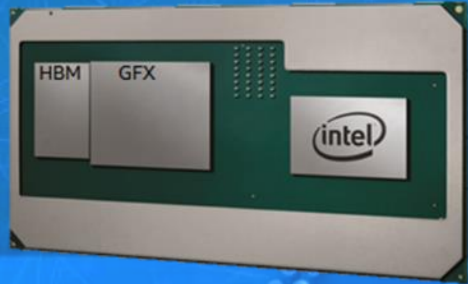




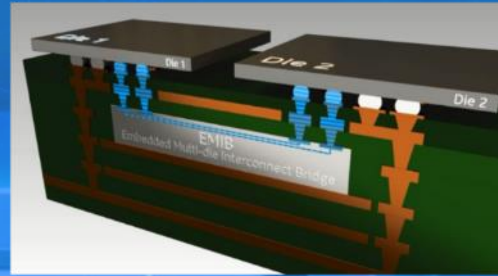


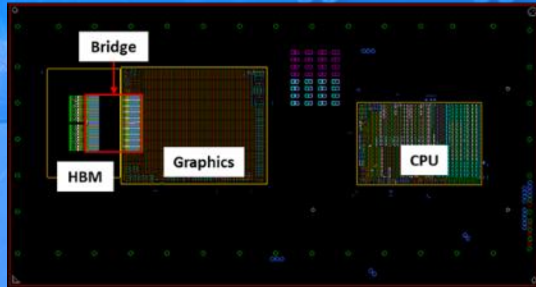
Our Focus : Push Wiring Density for Increased BW + Improved Power Efficiency

In the 2D Space….Intel has EMIB*



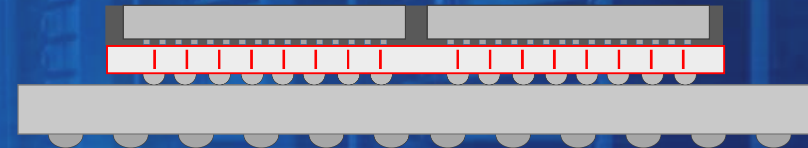
EMIB





- **Localized** high density wiring
- No practical limits to die size
- Standard assembly process
- Bridge manufacturing much simpler
- Bridge silicon costs < Silicon interposer - No TSVs, Significantly less silicon area

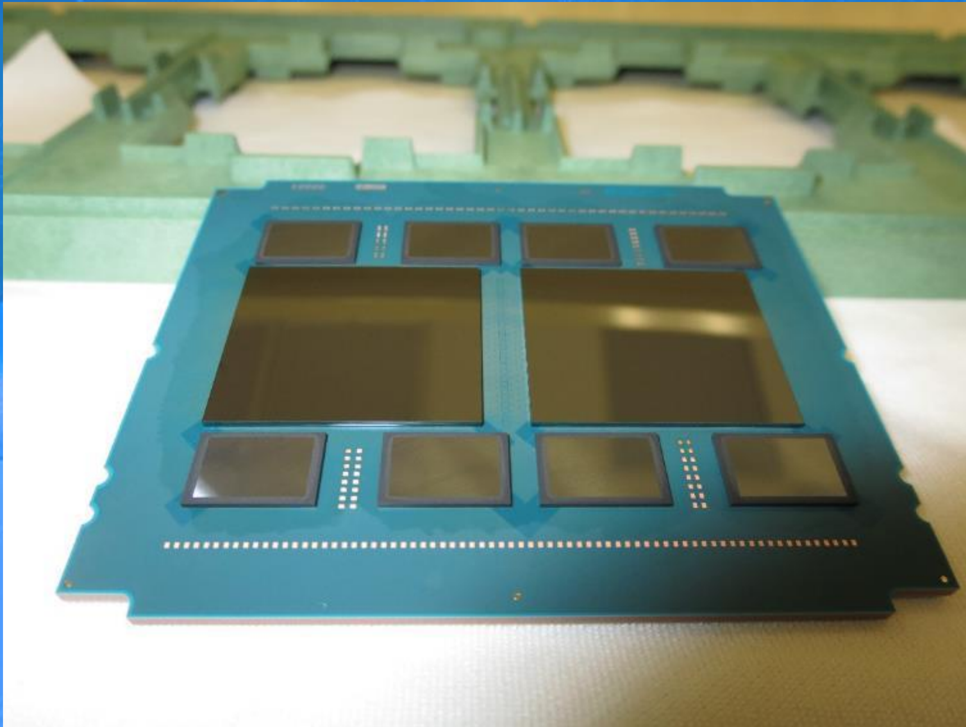
Silicon Interposer

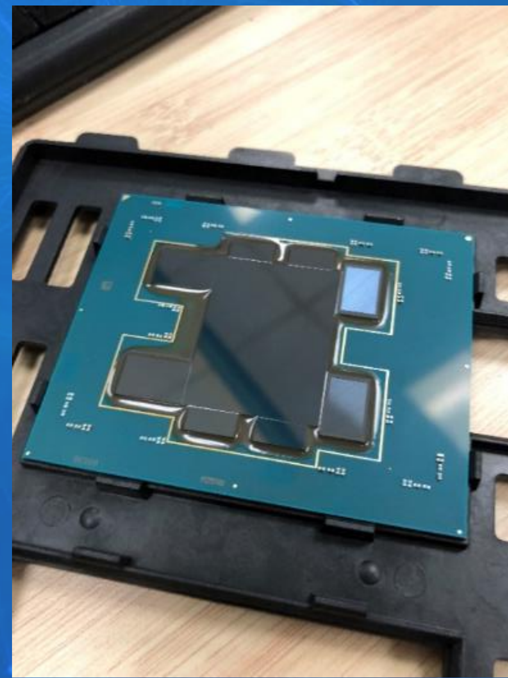


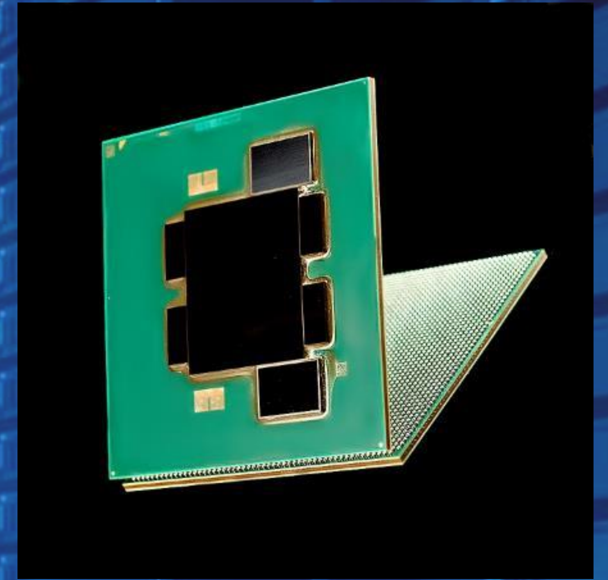
- CTE Matched with Si : Low stress on low-K ILD
- Excellent Chip-Attach Alignment
- Pitch scaling
- Interposer size is typically limited by reticle field : Active Efforts in place to develop larger than reticle interposers
- TSV capacitance impacts signal integrity of off-package links
- Interposer attach adds an extra chip attach step

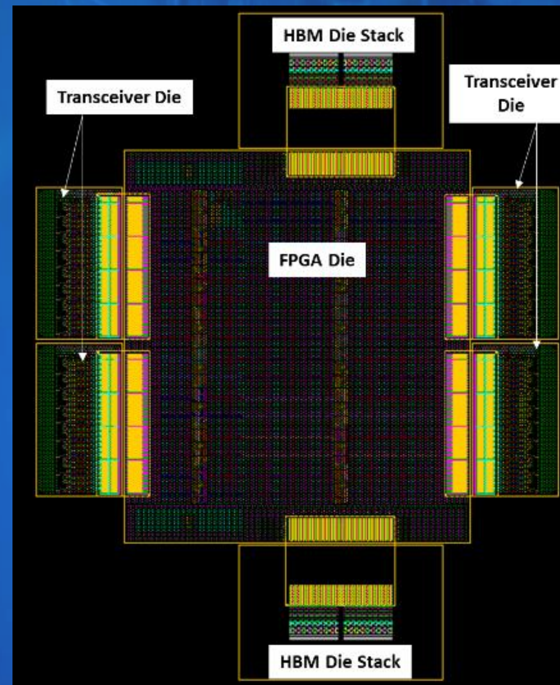
*Embedded Multi-Die Interconnect Bridge

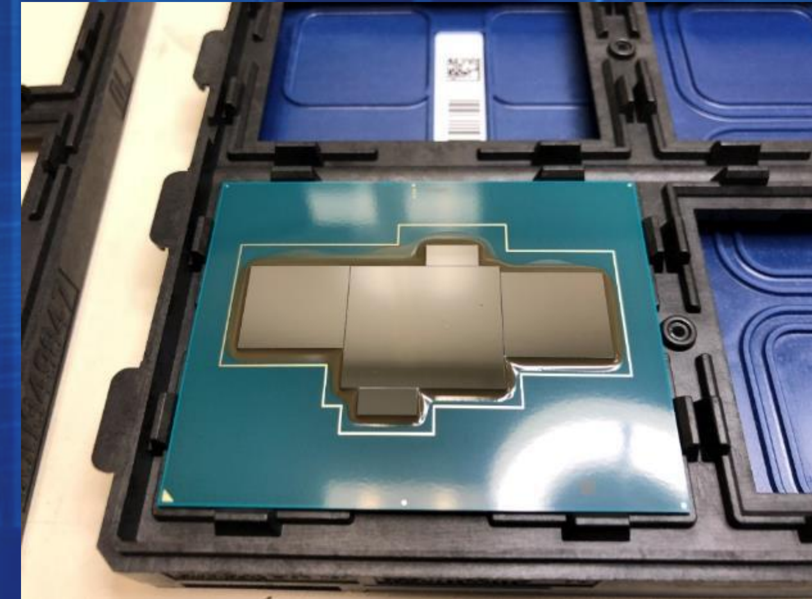
EMIB Offers Flexibility





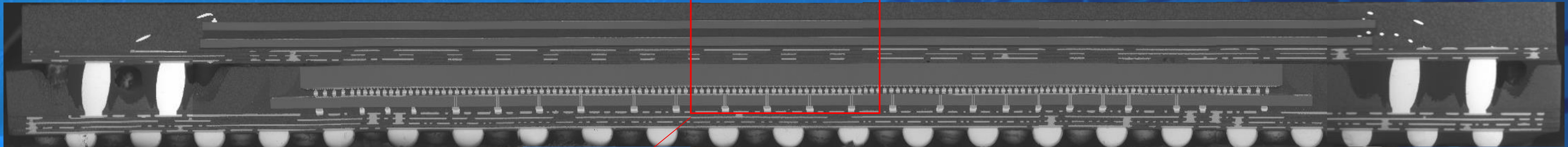


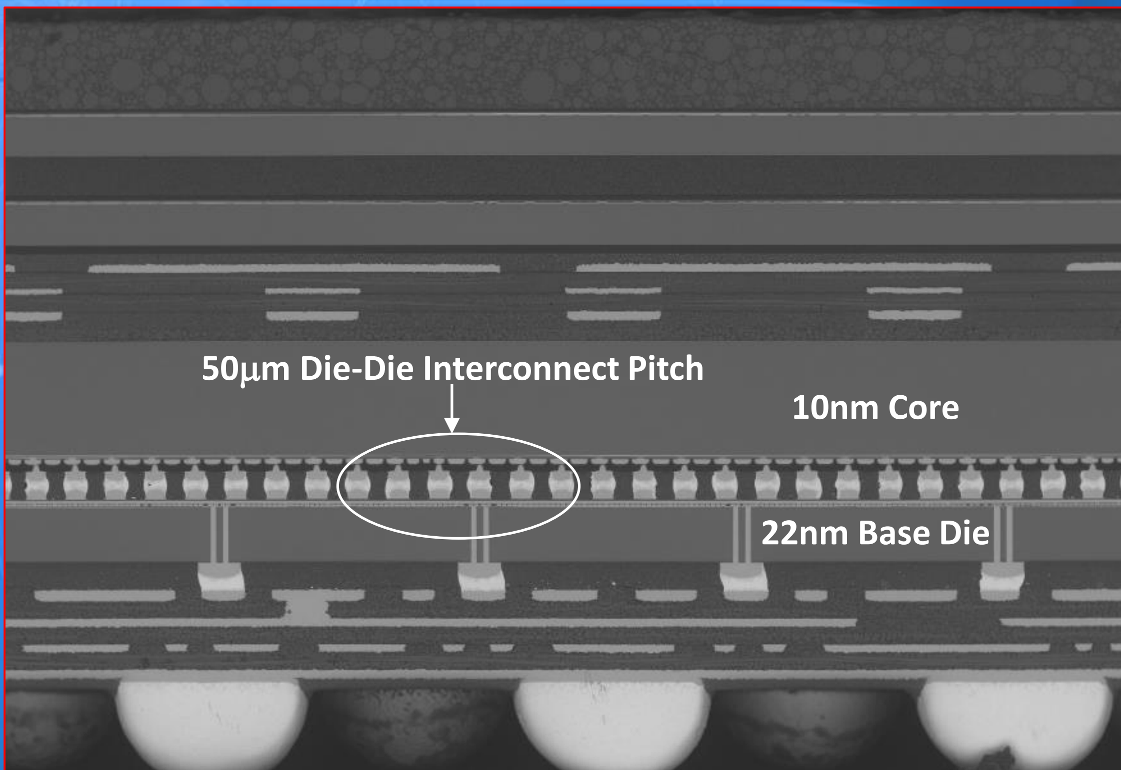




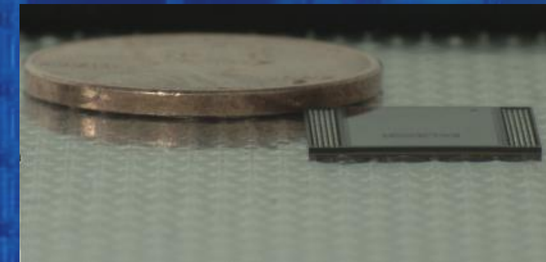
- In High Volume Production
- Multiple Bridges, Multiple Bridge Sizes and Bridge Technologies possible
- Die from Different Foundries
- Large Overall Die Area enabled

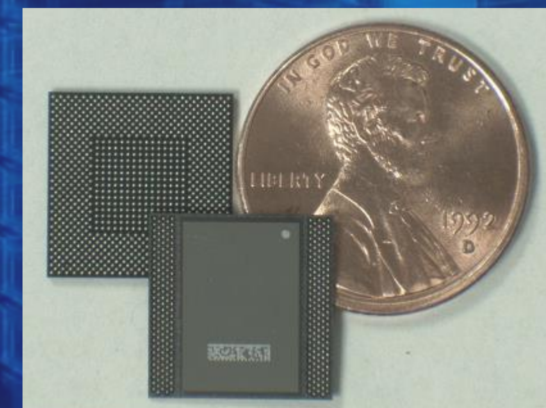
In the 3D Space….Intel Has Foveros



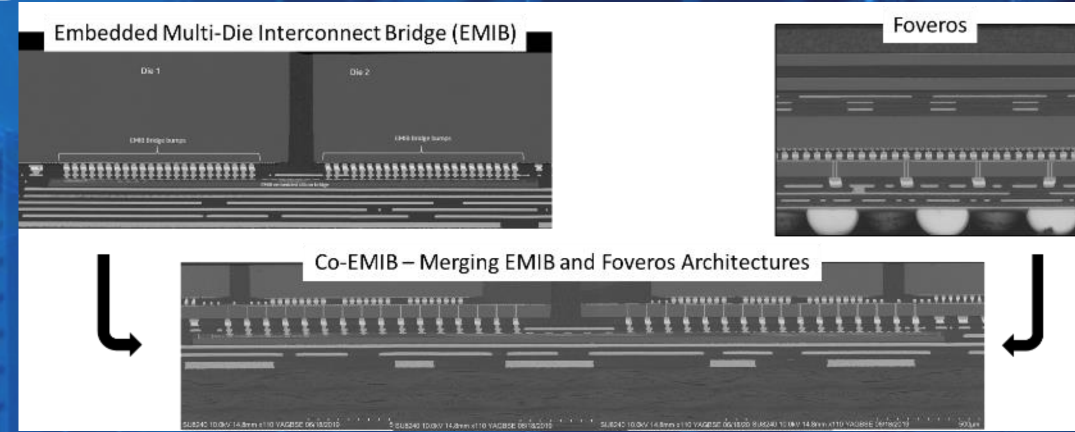


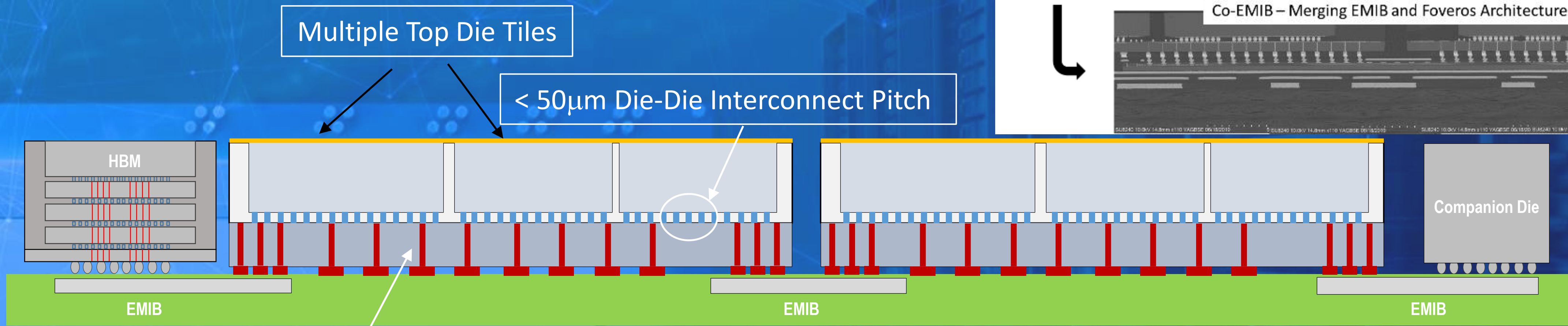
- Stacking allows functional partitioning, Smaller footprint and improved interconnect density



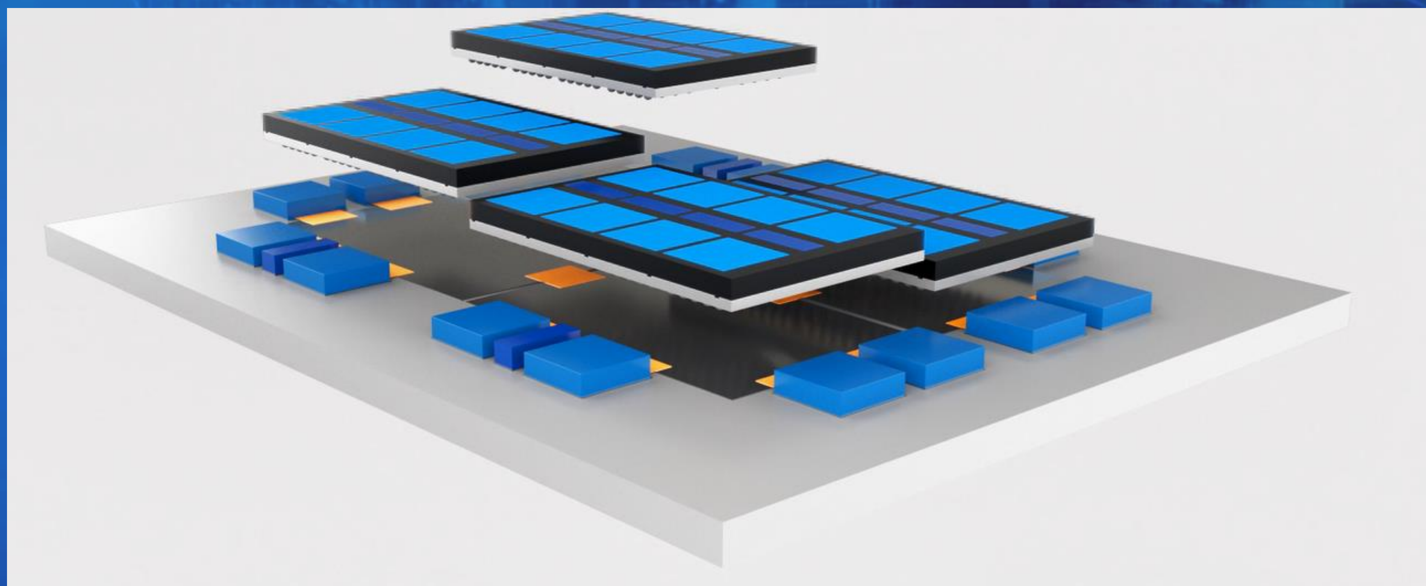


Co-EMIB: Blending 2D and 3D

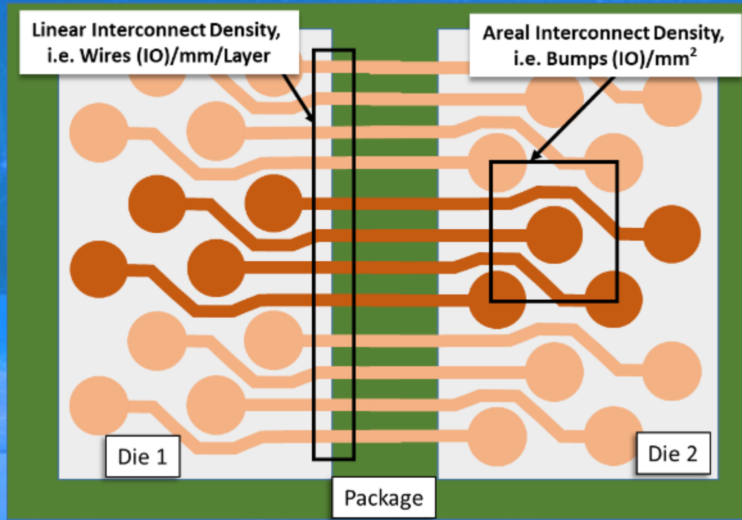




- Architecture enables >> reticle sized base die & High Density Bridge links to companion Die
- Increased Partitioning Opportunities



3D MCP Landscape



Bump/mm^2

100000
10000
1000
100

0
10
20
30
40
50
60

Bump Pitch (μm)

Cu-Cu Interconnects

Bonded Structure
Thru-Si Via (TSV)
Top Thin Wafer
Bottom Wafer
5 μm

IEEE ELECTRON DEVICE LETTERS, VOL. 27, NO. 5, MAY 2006

335

Three-Dimensional Wafer Stacking Via Cu–Cu Bonding Integrated With 65-nm Strained-Si/Low-k CMOS Technology

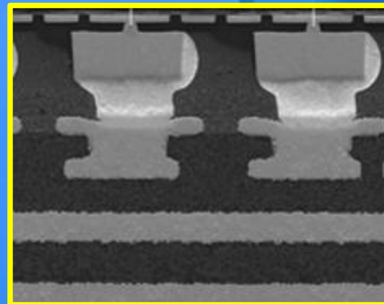
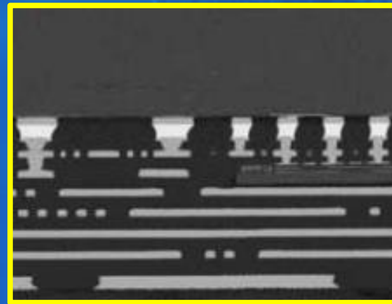
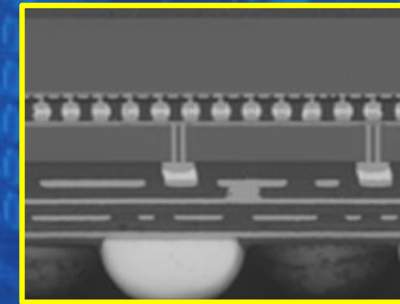
P. R. Morrow, C.-M. Park, S. Ramanathan, M. J. Kobrinsky, and M. Harmes

Solder Based Interconnects

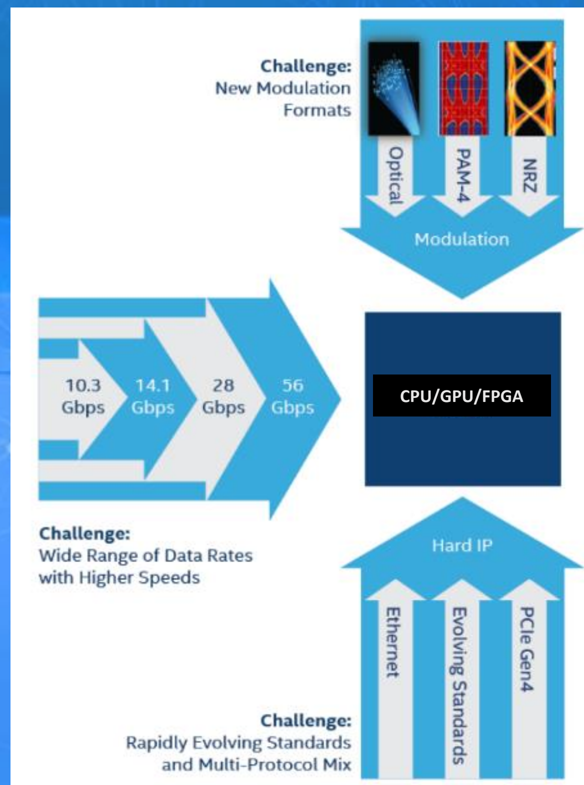
Somewhere Below 25μm Bump Pitch, the Transition from Solder Based Interconnects to Cu-Cu interconnects will be needed

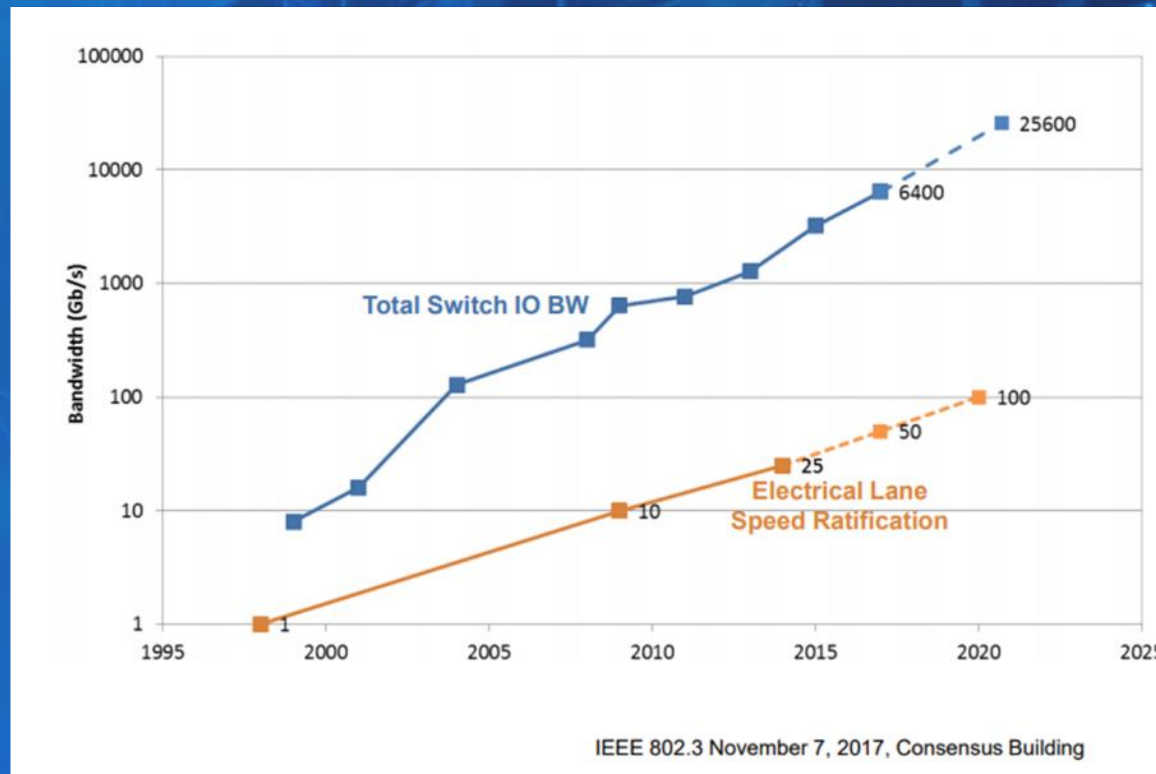
Intel's Package Interconnect Roadmap

	Traditional Flip-Chip	2D - EMIB	3D - Foveros
Today	100µm	55µm	50µm
Future	90µm	30µm-45µm	20µm-35µm (Solder) <20µm (Non-Solder)

Performance Challenges : Off-Package High Speed IO

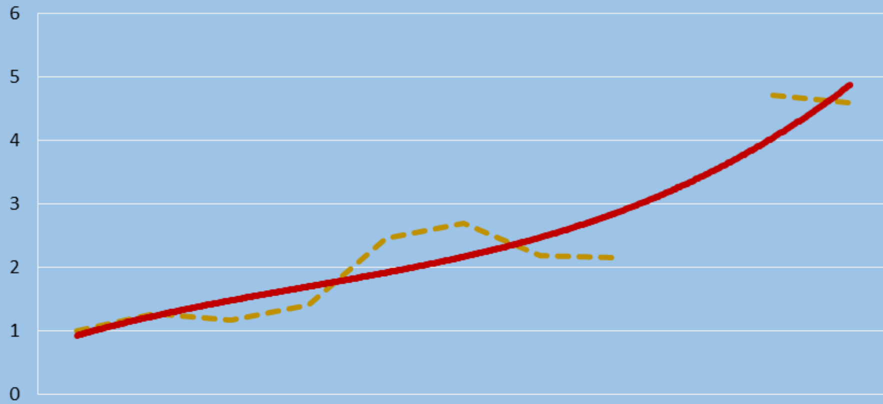




High Bandwidth, Off-Package Electrical & Optical Signaling Technologies Needed

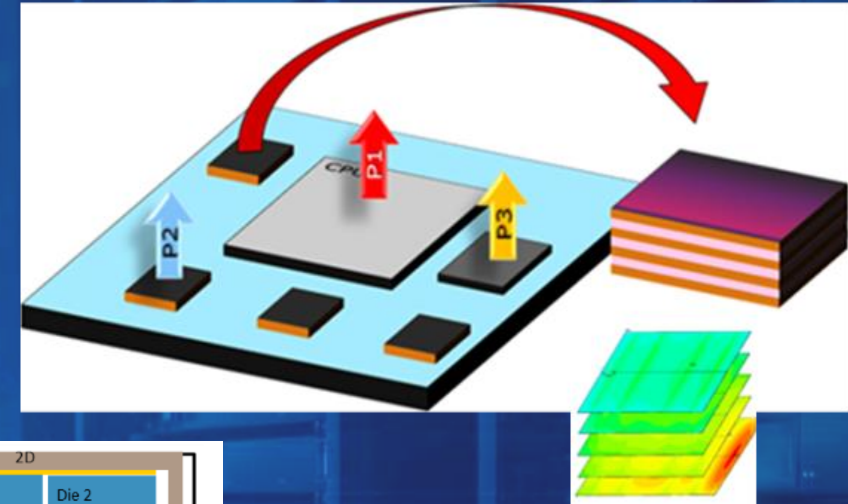
Performance Challenges : Thermals

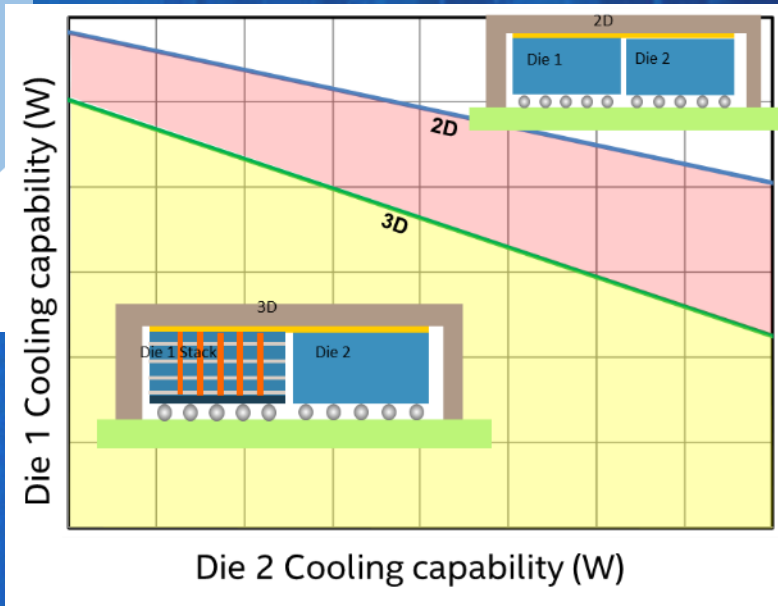
Normalized Power Density Trends



Si node scaling

Local hotspot density increasing gen-o-gen; likely to hit thermal wall ahead of power delivery wall





Key challenges are Raw power, Power density and Thermal Cross-talk.
Thermal Co-design, Improved TIMs + Interface Control Needed

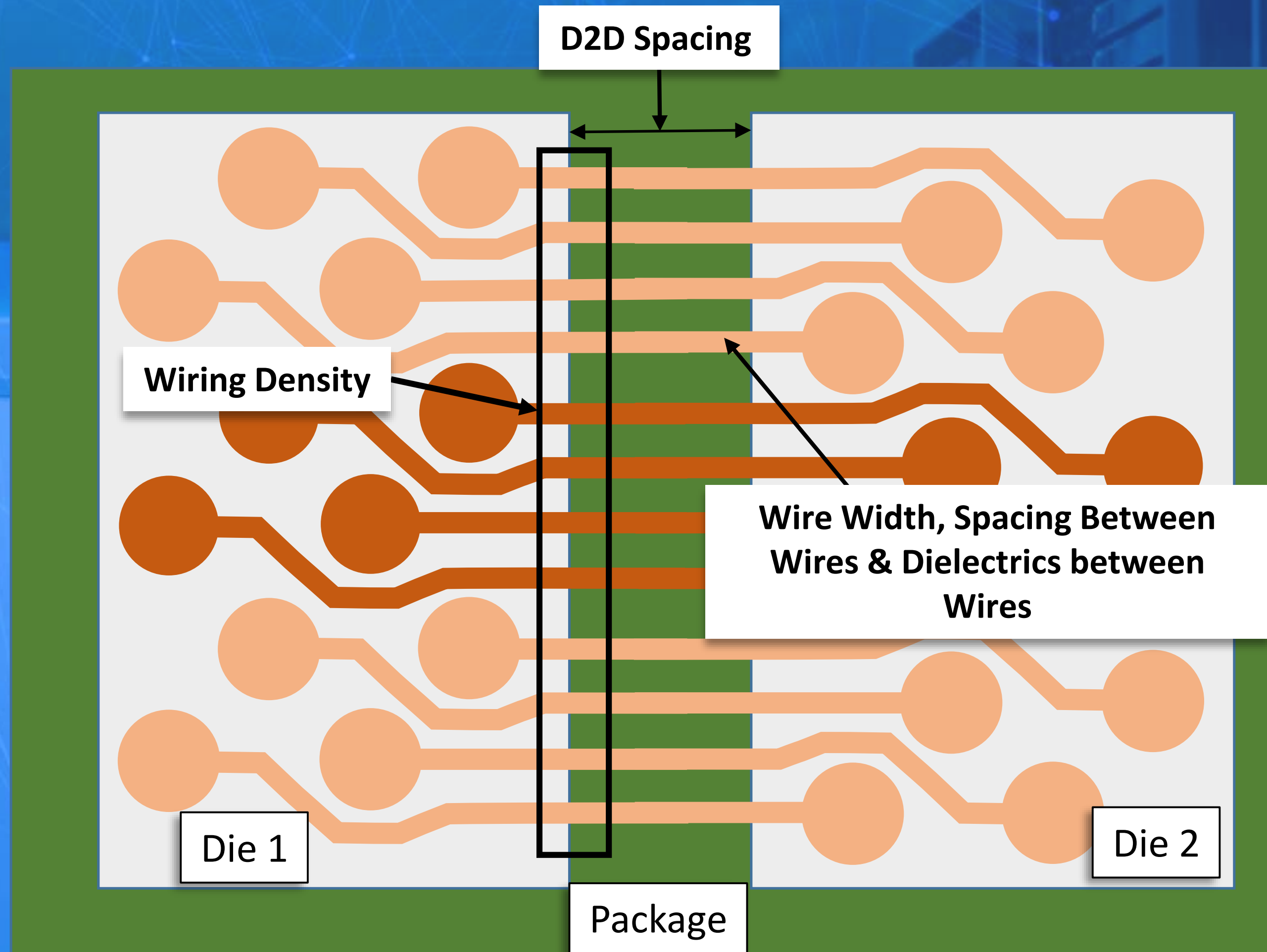
Key Takeaways

- Intel has a long and storied history in driving innovations in packaging
- We have always seen the package as a compact HI platform
- Interest in high bandwidth, low power die-die links drives the evolution of high density interconnects in 2D and 3D architectures
- Intel has a comprehensive current packaging portfolio and a forward looking roadmap

Back-up

Signaling Performance in Dense Interconnects

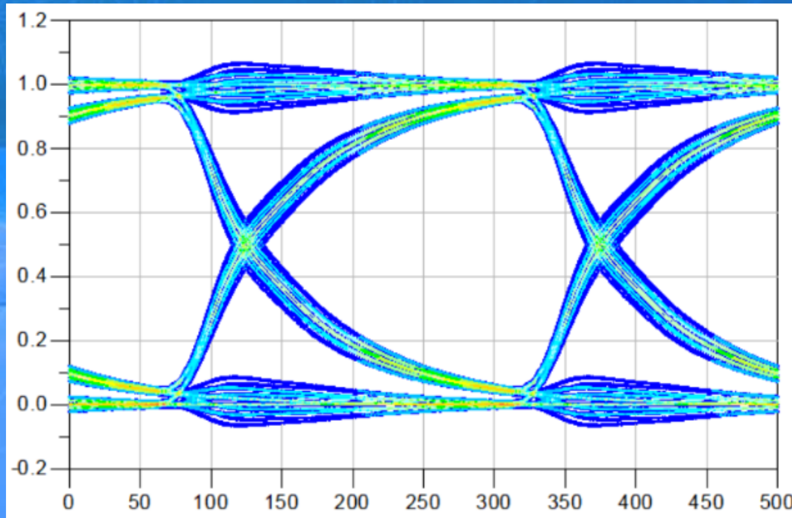
- Key factors Affecting Signal Integrity and Power efficiency



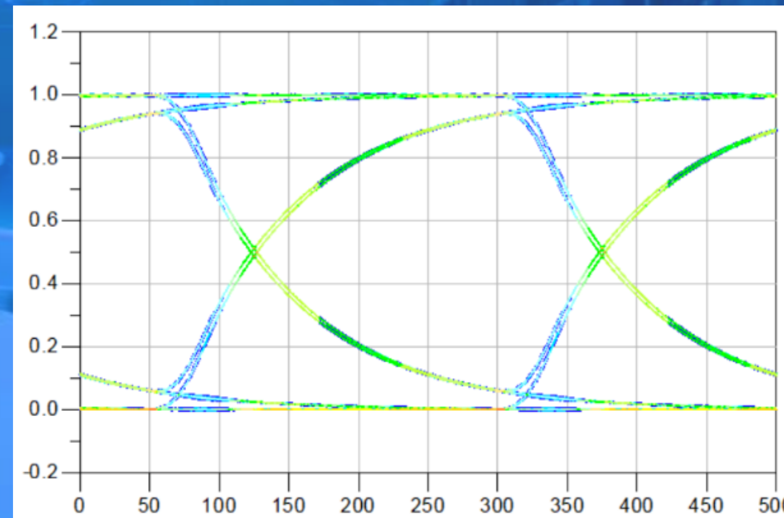
Signaling Performance in Dense Interconnects

- Wire Widths, Spacings and Signal & Ground Ratios are key knobs for improving Signal Integrity

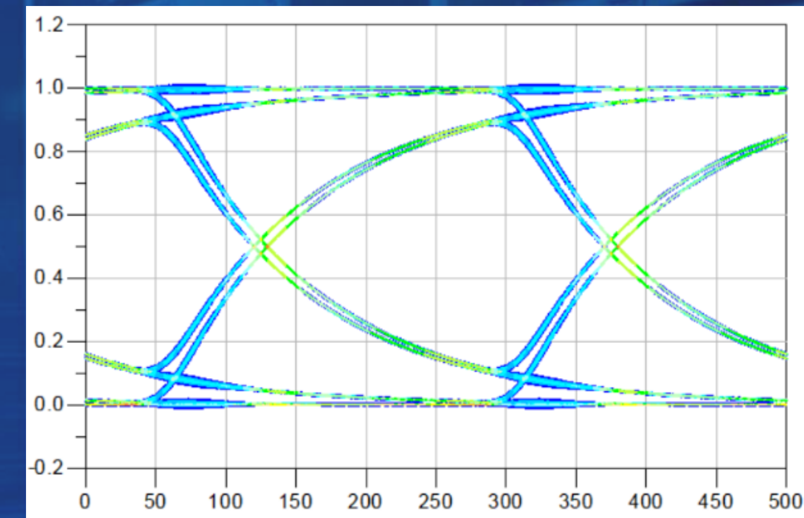
3S: 1G 500 Wires/mm

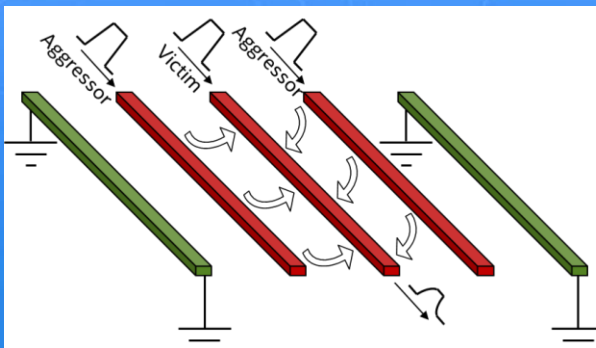


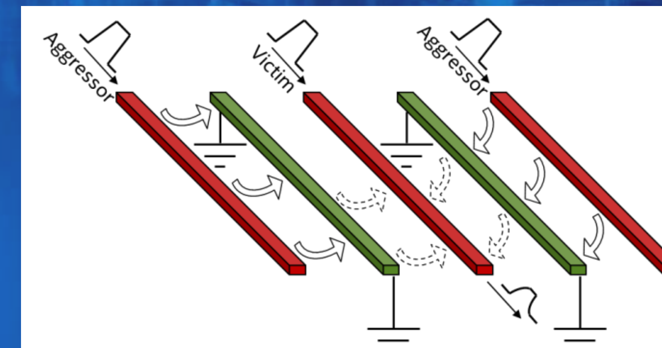
1S:1G 500 Wires/mm



1S:1G 1000 Wires/mm







1. For Details on Optimization See A.C. Durgun, Z. Qian, K. Aygün, R. Mahajan, T.T. Hoang, S. Shumarayev, "Electrical performance limits of fine-pitch interconnects for heterogeneous integration," presented at 2019 IEEE 69th Electronic Components and Technology Conference, Las Vegas, NV, 2019
2. Simulations above assume a link length of 0.8mm