

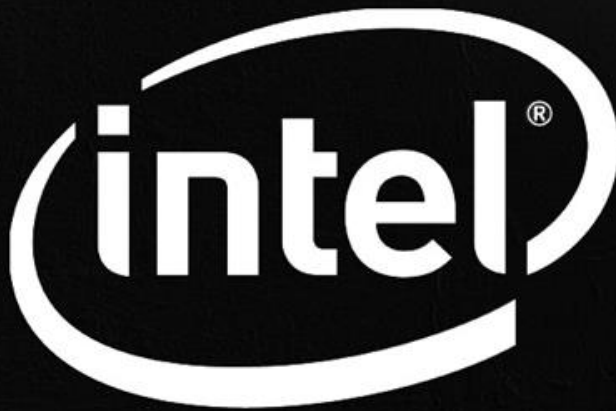
ADVANCED PACKAGING ARCHITECTURES FOR HETEROGENEOUS INTEGRATION



RAVI MAHAJAN

INTEL FELLOW, ASSEMBLY AND TEST TECHNOLOGY DEVELOPMENT

PWRPACK, NOVEMBER 1, 2019



Legal Notices

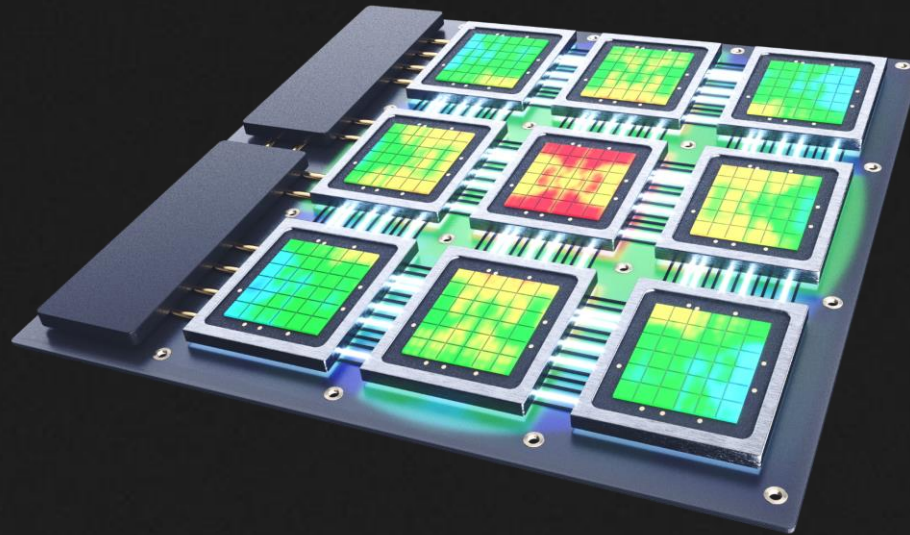
This presentation contains information provided by Intel Corporation ("Intel"), and may refer to Intel's plans and expectations for the future, which are provided for discussion purposes only and are subject to change without notice. Forward-looking statements involve a number of risks and uncertainties: 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.

Intel's Packaging Vision

Raja Koduri (2018)

Develop and own leadership technology to connect chips and chiplets in a package to match the functionality of a monolithic SOC

High density interconnect that enables high bandwidth at low power is essential to realize this vision

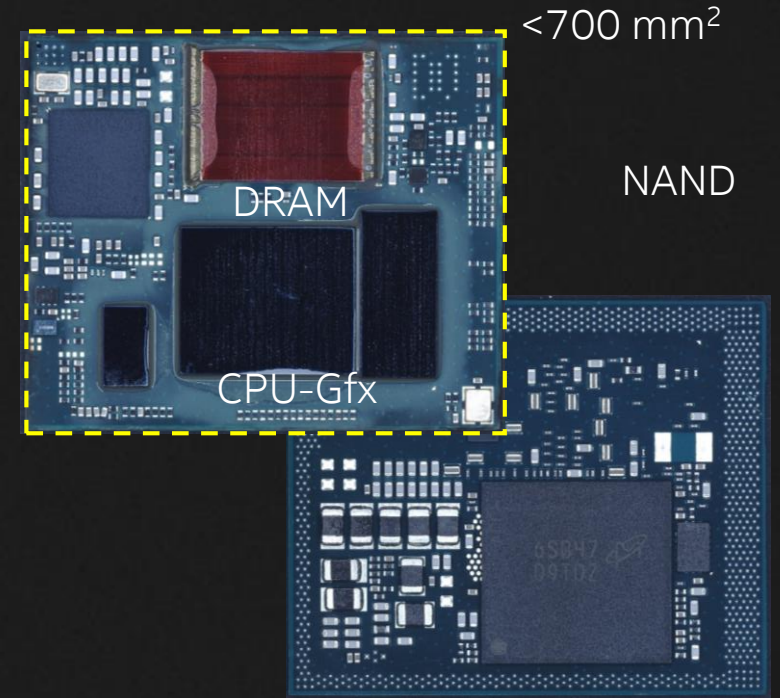
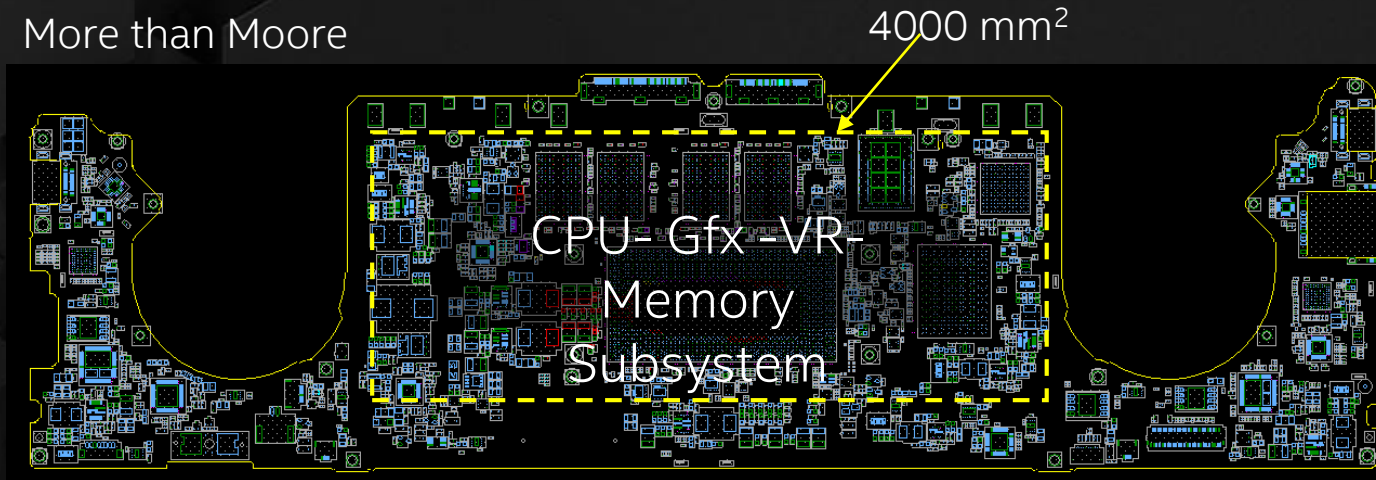


PACKAGE TECHNOLOGY FOCUS

- Thin/small footprint client packages
- High speed signaling
- Interconnect scaling – density and pitch

INTEGRATION – SIZE MATTERS

More than Moore



PCB Integration

- Limited Interconnect Density → Limited BW
- Long Interconnects → Increased Power
- Large Form Factor

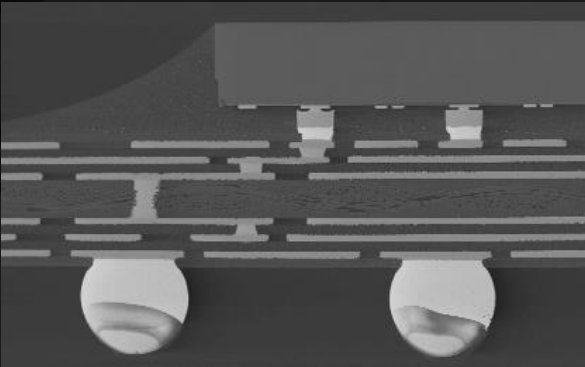
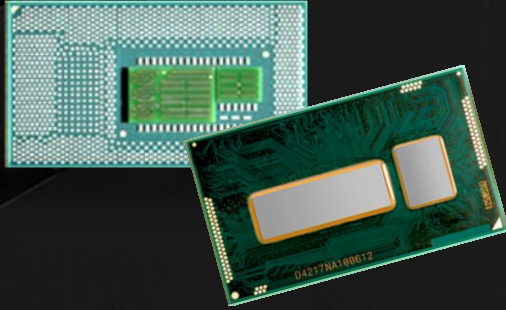
Heterogeneous Packaging

- Smaller system footprint
- Better VR Efficiency/Loadline
- High speed Signaling (Ghz)
- Improved data Latency (ps)
- Mixed Node Integration

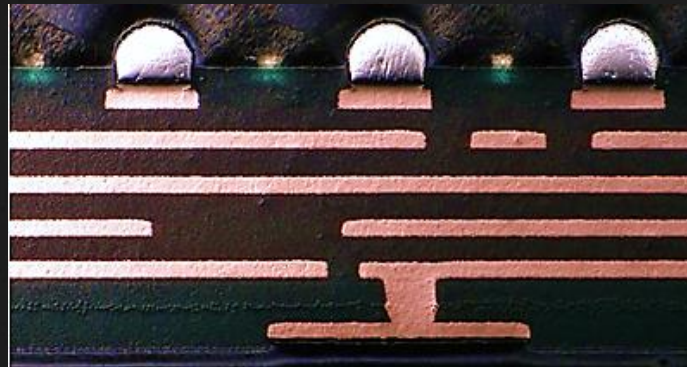
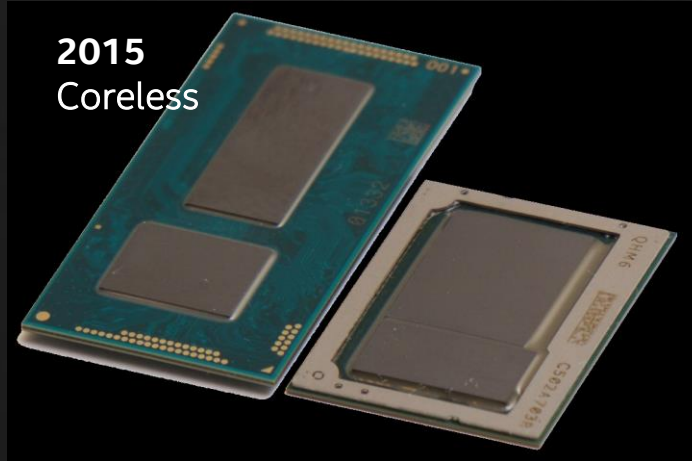
Smaller is better

ULTRA THIN PACKAGING

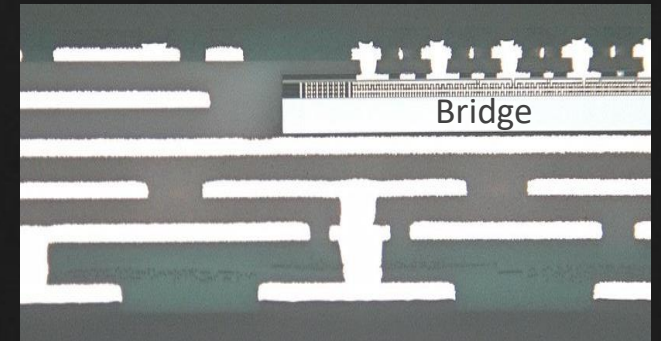
2014
100 μ m Core thickness



2015
Coreless

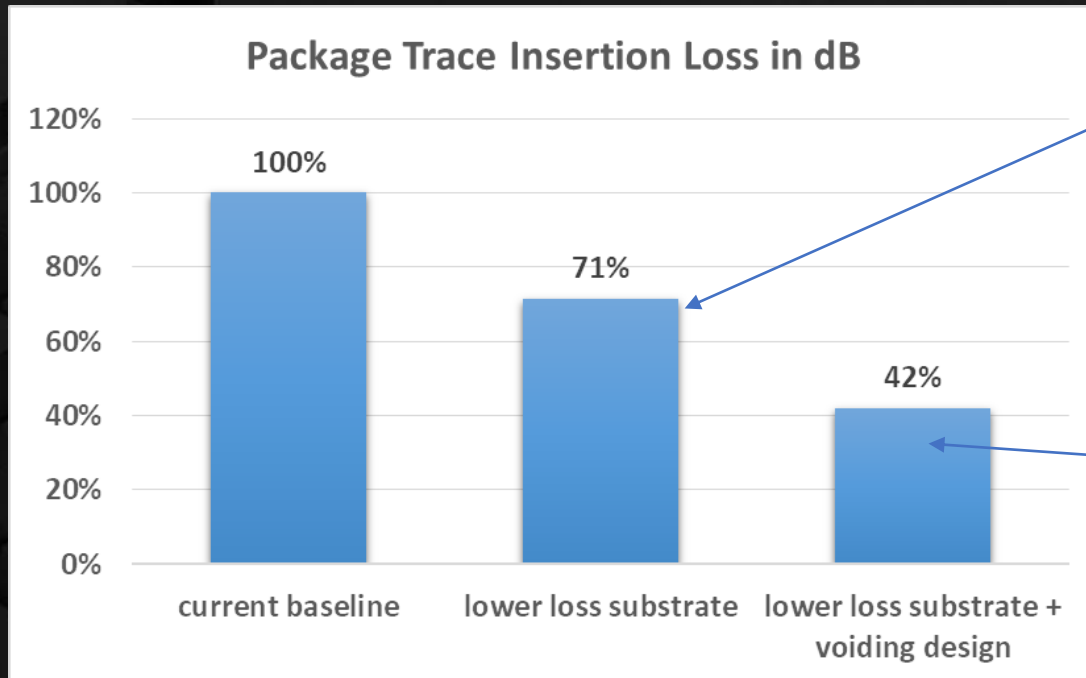


Future
Coreless with Embedded Bridge



Ultra thin Core and Coreless Package for Thin Client Applications Is an Important Enabler

HIGH SPEED SIGNALING



Conventional Routing



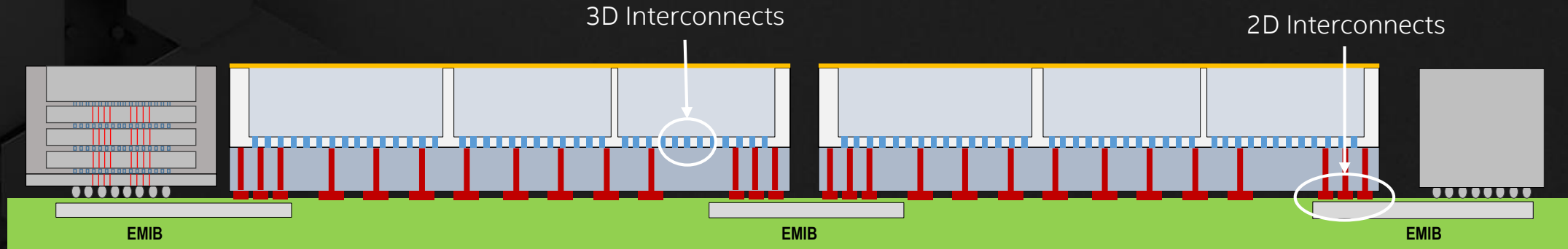
Routing with Voiding



Enabled 112 Gbps; Working towards 224 Gbps

Lower loss through Dielectric Material inventions and metal surface roughness
Design IP with routing/ plane templates and dielectric stacks

HIGH DENSITY, HIGH BANDWIDTH INTERCONNECTS



High Bandwidth, Low Power, “Wide & Slow” Parallel Links drive Need for High Density Die-Die Interconnects

DARPA 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	High
Energy Efficiency (<1000um)	0.1pJ/bit	1-10pJ/bit
Throughput per pin	2Gbps	30Gbps
Packaging Complexity	High	Low
Packaging Cost	High	Low

Physics

Economics?

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)

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

GDDR6 Memory System: Four 16Gbps x32 GDDR6 DRAMs

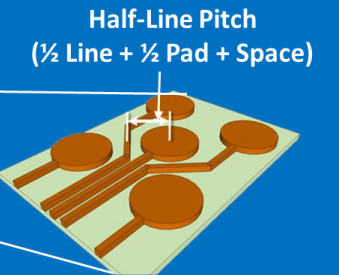
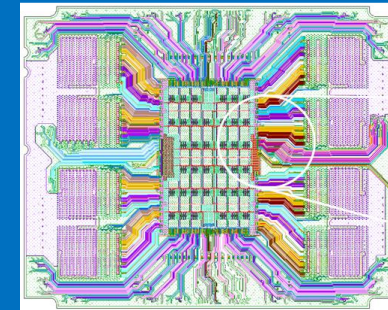
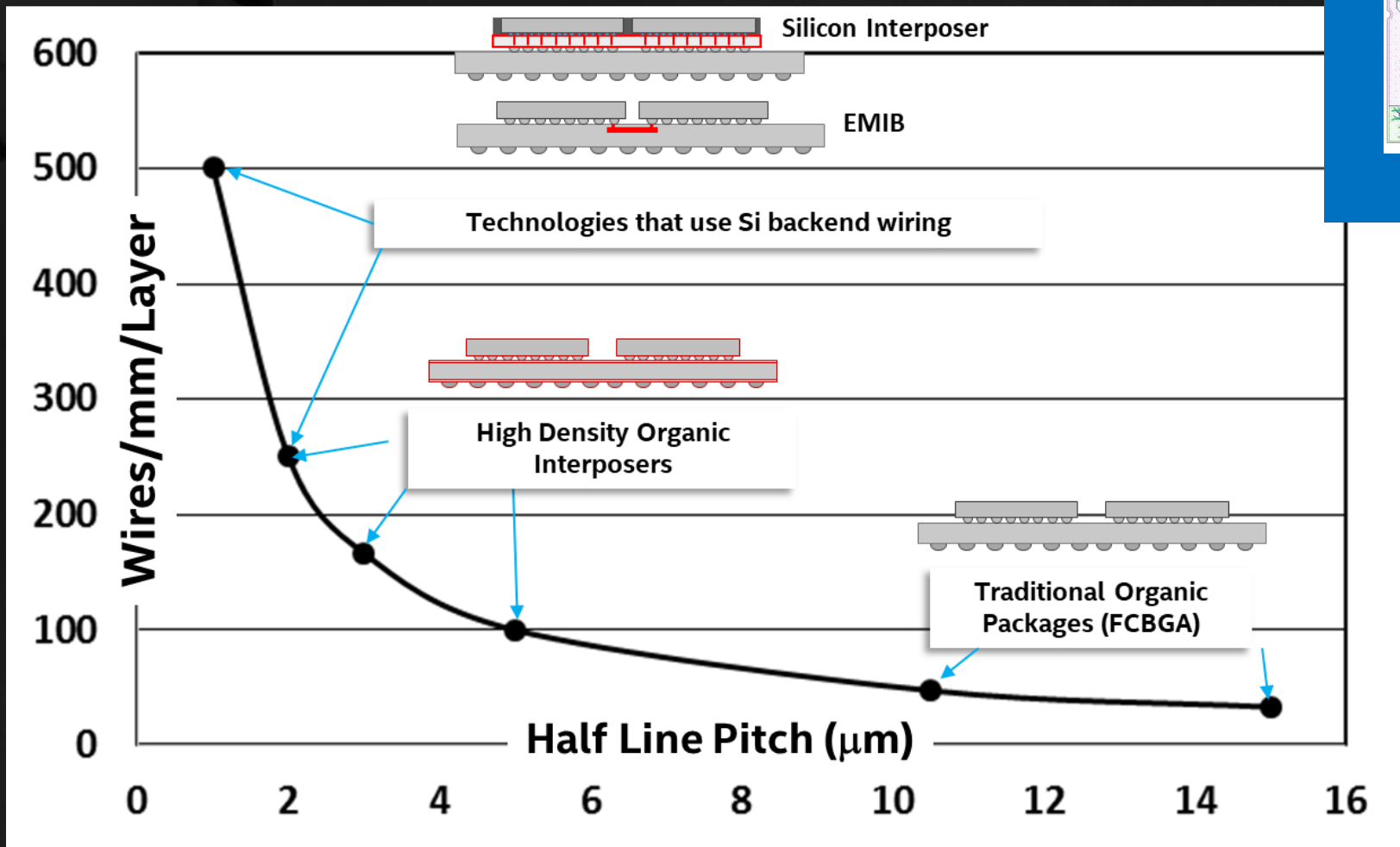
HBM2 Memory System: Single 2Gbps HBM2 Device

	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

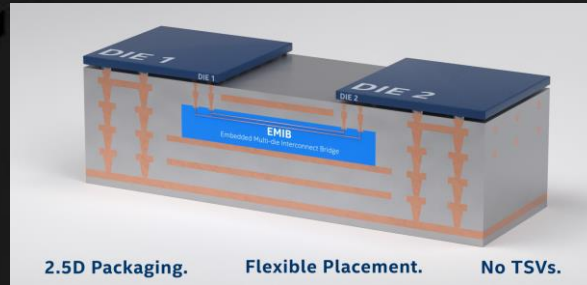
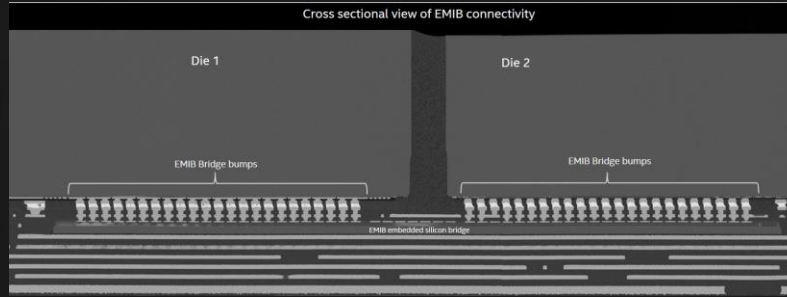
[1] Source: Rambus Inc.

[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>

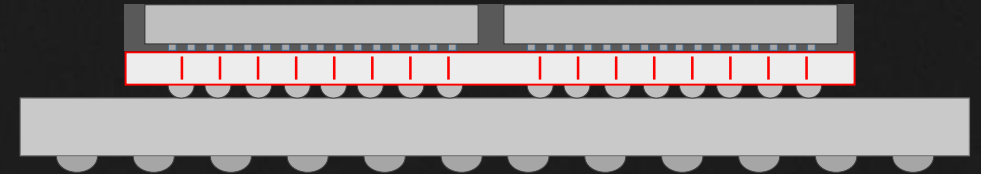
2D MCP LANDSCAPE



INTEL EMIB TECHNOLOGY : HIGH DENSITY 2D



Silicon Interposer



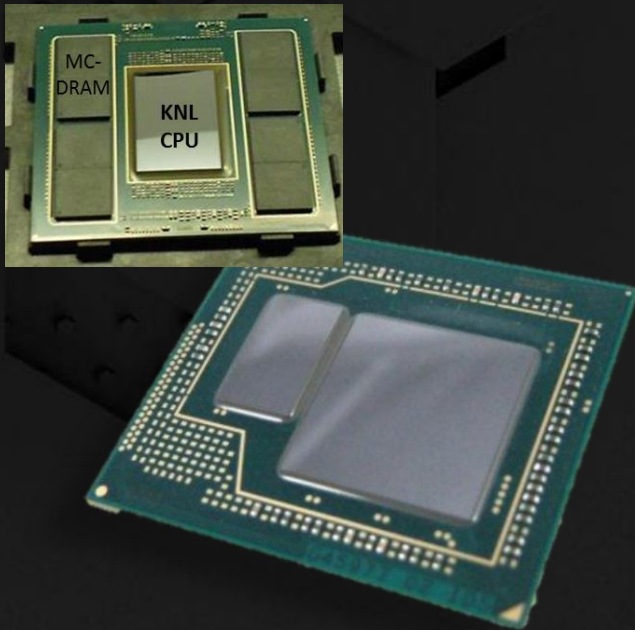
- Localized high density wiring
- No practical limits to die size
- Flexible : Allows Bridge Mix and Match
- Standard assembly process
- Bridge manufacturing much simpler
- Bridge silicon costs < Silicon interposer – No TSVs, Significantly less silicon area

- 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 off-package signal integrity
- Interposer attach adds an extra chip attach step

DENSITY SCALING (PLANAR)

FCBGA

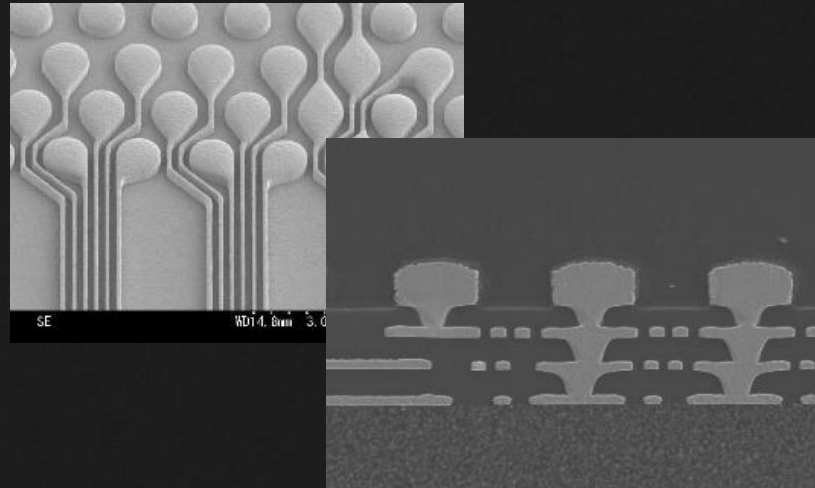
Typical Organic Package



IO/mm/lyr = 32 → 48

HYPER DENSITY FCBGA (WIP)

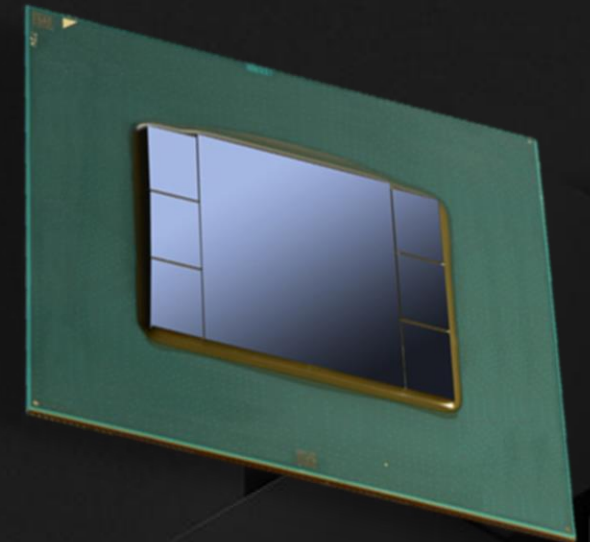
New Eqpt/ Litho based process on organics
or organic RDL for fine L/S, line vias



IO/mm/lyr = 64 → 250

EMIB

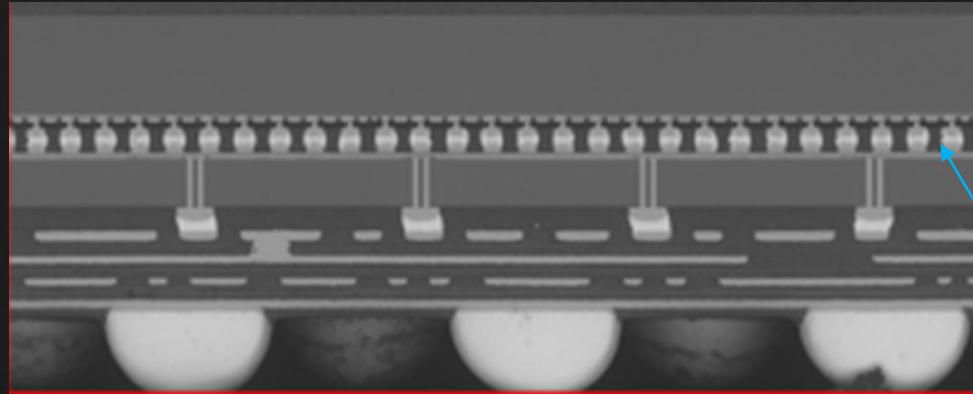
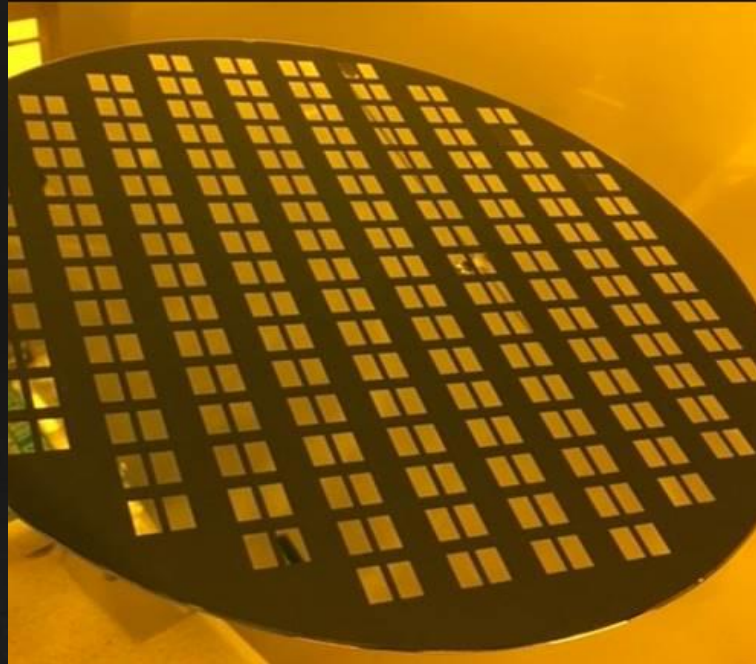
Embedding die with dense wiring in a
package cavity



IO/mm/lyr = 256 → 1000

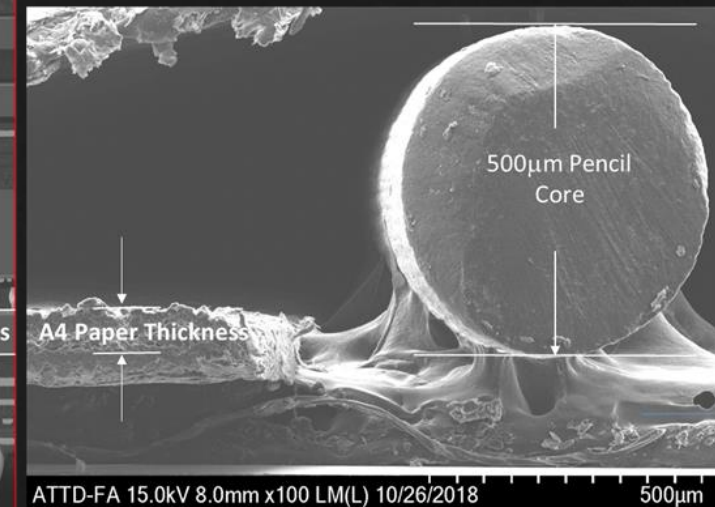
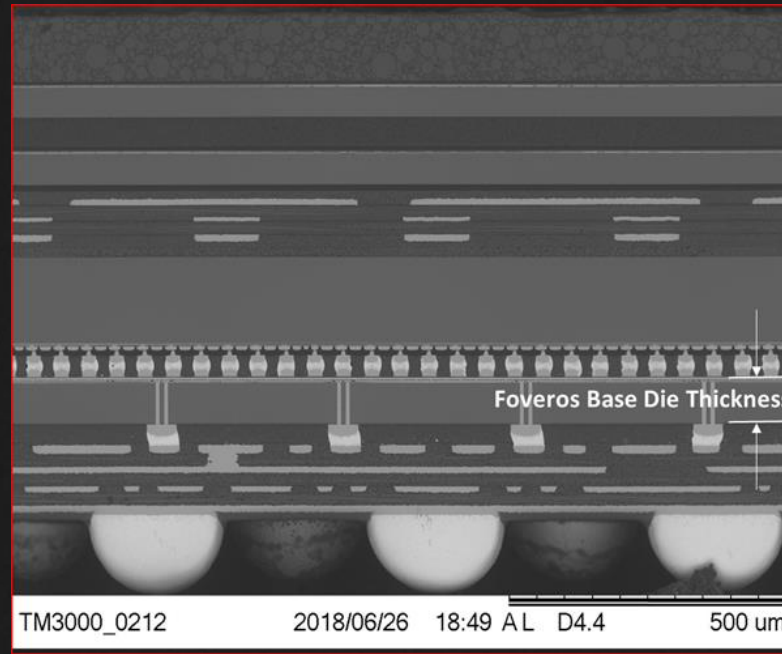
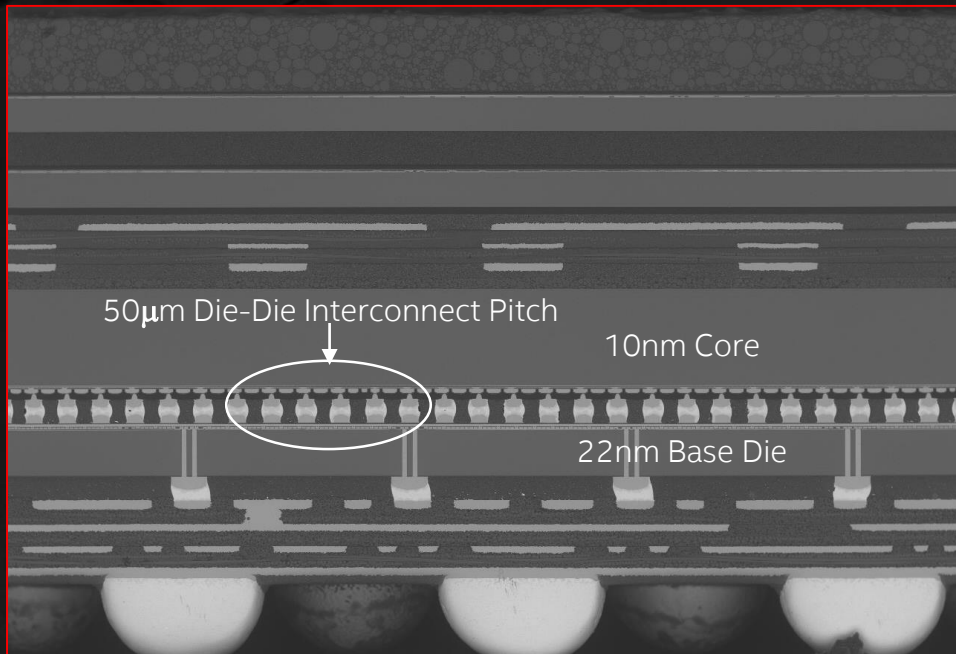
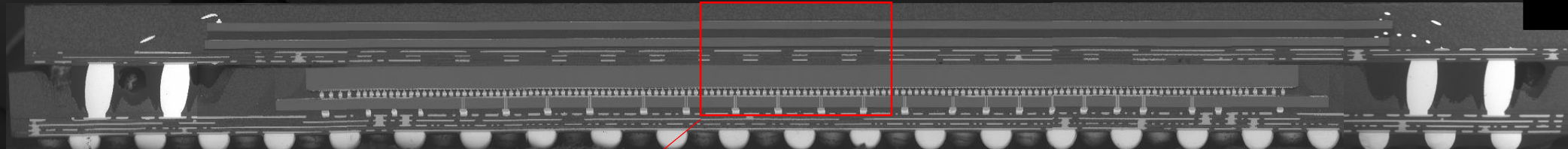
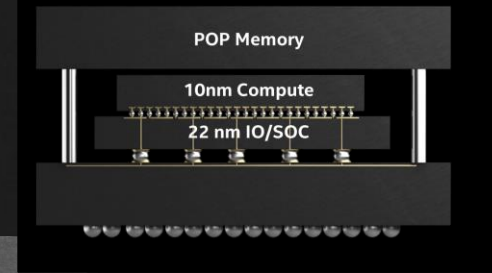
DENSITY SCALING (3D)

LKF-2018



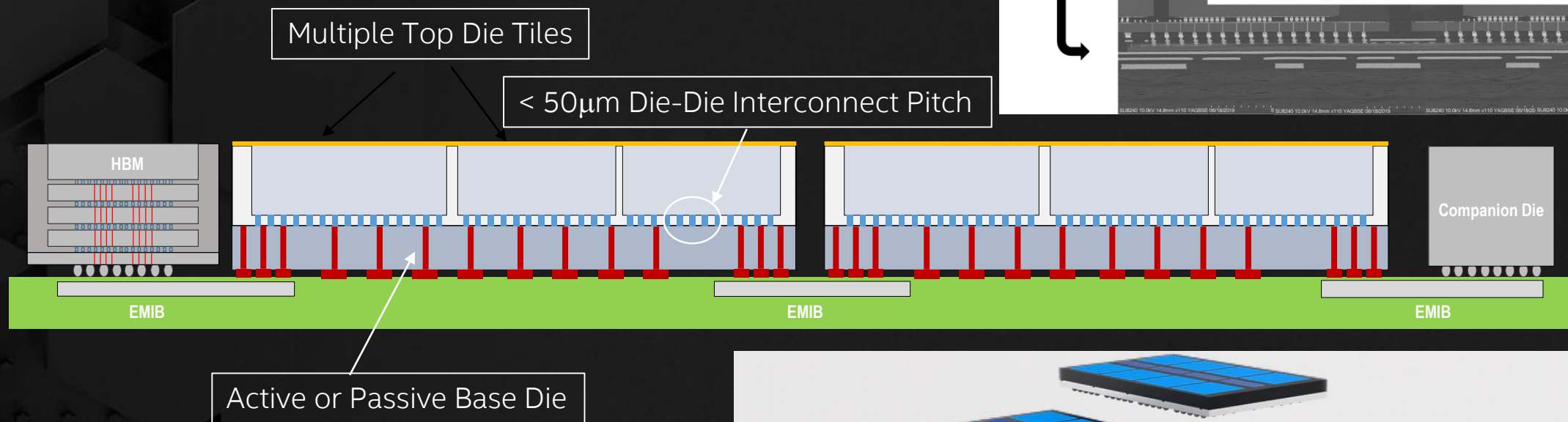
Pitch = $50\mu\text{m}$ to $10\mu\text{m}$
 $\text{IO}/\text{mm}^2 = 400$ to $10,000$

INTEL FOVEROS TECHNOLOGY : HIGH DENSITY 3D

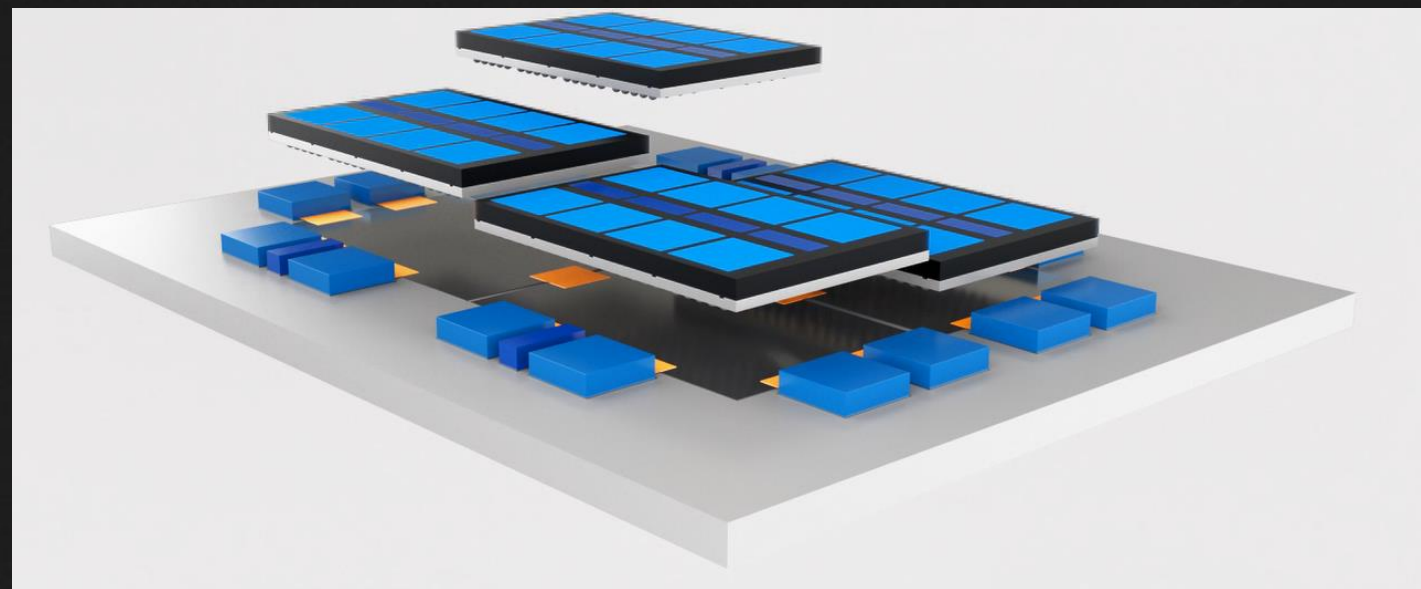
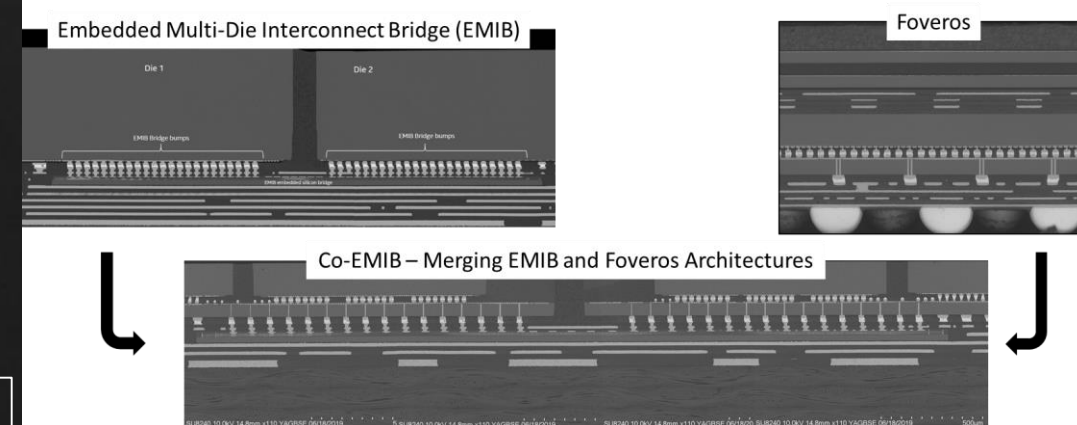


Single Tile on Base Die @ 50 μ m Pitch in Advanced Development
TD Focus : Multiple Tiles @ Reduced Pitch

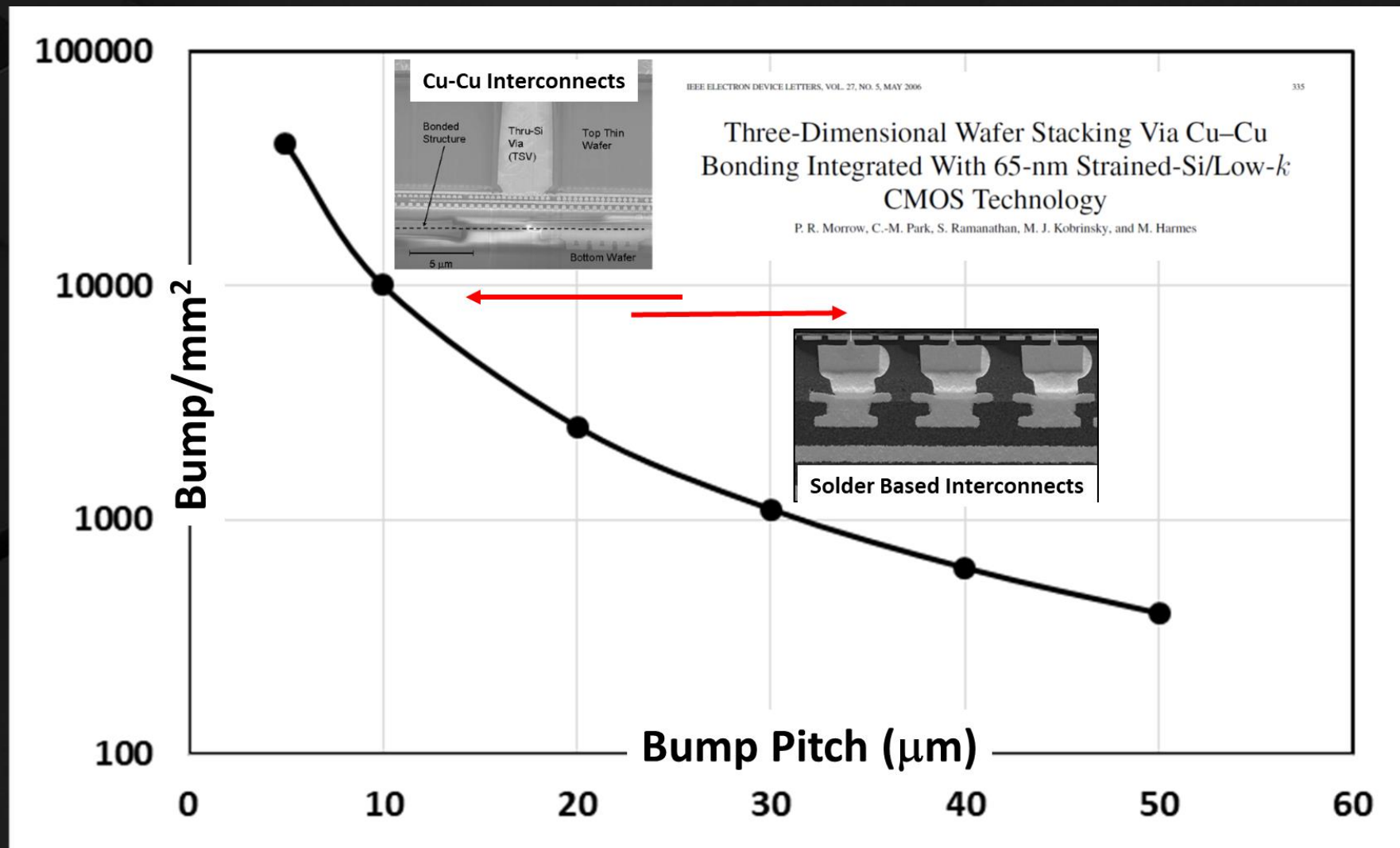
CO-EMIB: BLENDING 2D AND 3D



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



3D MCP LANDSCAPE



KEY TAKEAWAYS

- Interest in Advanced MCP Architectures driven by need for high bandwidth, low power IO Links
- EMIB, Foveros and Co-EMIB are key Dense MCP Building Block Technologies
- MDIO & AIB are Leadership IO Technologies that are Co-designed using our Dense MCP technology Portfolio
- Together they give us architectural capabilities to Scale-up and Scale-out Heterogeneous Compute Elements

