Too Many Package Options: What Makes Sense for Your Application?

- TRACK INNOVATION
- IDENTIFY TRENDS
- ANALYZE GROWTH
- INFLUENCE DECISIONS

RELEVANT, ACCURATE, TIMELY

E. Jan Vardaman, President and Founder

TechSearch INTERNATIONAL

Unprecedented Era of Change for Packaging

- **Economic advantage of silicon scaling is gone**
  - High cost of moving to next silicon node
  - High cost of fabrication includes design, mask, and fab process
  - Only a limited number of foundries can afford to participate for the limited number of companies at advanced nodes
- **Heterogeneous integration provides an opportunity to achieve economic advantages lost with end of pure silicon scaling**
  - Many options for the package including silicon interposers, FO on substrate, chiplets, and variations of 3D stacking
- **Heterogeneous integration offers improved SI, PI, lower inductance and thermal resistance, form factor advantages**
  - Co-design of silicon and package essential
  - Thermal issues must be addressed
  - Material selection is important

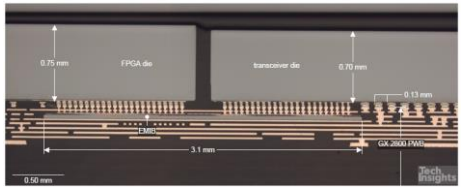
High-Performance Packaging: Drives the New Era

- **Markets**
  - Networking, data centers, AI/accelerators, machine learning, gaming, 5G infrastructure, edge computing
- **Key performance metrics driving the adoption of Si interposer, 3D, and HDFO**
  - Low latency
  - Bandwidth & data rate increase (GHz, Gbps)
  - Better power efficiency (pJ/bit) and improved power delivery
  - Routing density increase (# of lanes per mm/layer)
  - IO density (IO/mm^2)
- **Cost Drivers**
  - Si partitioning in advanced nodes
  - Reuse of IP for time to market & reduction in cycle time

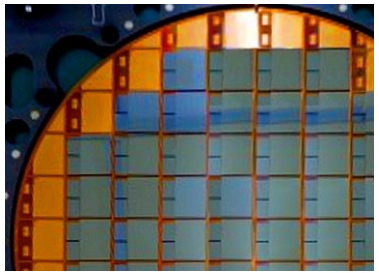
Many Different Packaging Options

- **Silicon Interposers**
  - Xilinx (many FPGA products) started in 2012, AMD shipped in 2015 (GPU + HBM), NVIDIA (GPU + HBM), Google, Baidu, Broadcom, many others in production!
- **Fan-Out on Substrate**
  - Chip first (in production since 2016 at ASE) and chip last solutions (many designs)
  - In production for network switch at TSMC with InFO_oS (since 2018)
  - Amkor Substrate-SWIFT®, SPIL Fan-Out Embedded Bridge, TFME
- **Embedded bridge Intel, IBM, SPIL, TSMC, and others**
  - Intel's EMIB silicon bridge in an laminate substrate, embedded by substrate supplier (Intel Stratix 10 for AI accelerator)
  - Amkor, ASE, IBM, SPIL, TSMC offerings
- **RDL Interposers InFO-SoIS (System on Integrated Substrate), Samsung R-Cube™, Unimicron, and others**
  - Extend to finer feature (2µm L/S) switch to fab-like process
- **3D stacking**
  - Memory-on-logic (AMD's new 3D Vcache, Samsung's X-Cube) or logic-on-logic (Intel's Foveros)

Source: NVIDIA.

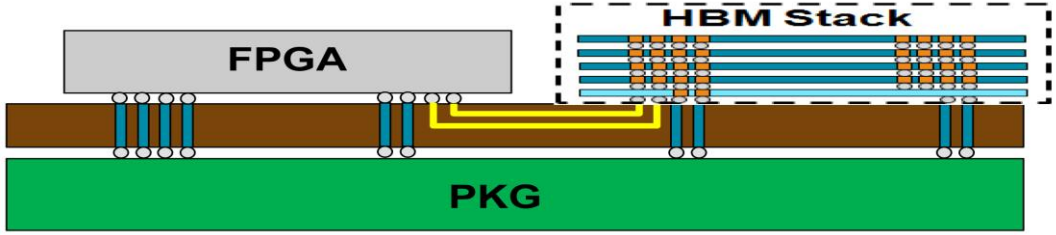


Source: Intel by TechInsights.



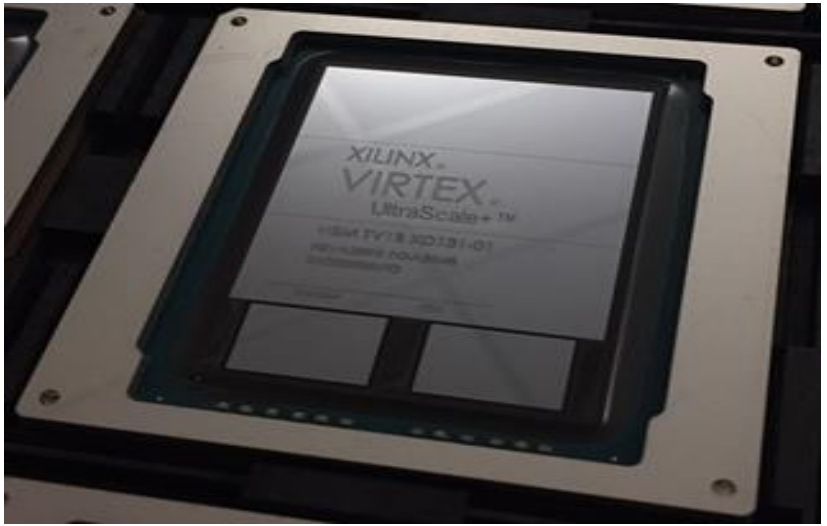
Source: ASE.

Why Si Interposer?



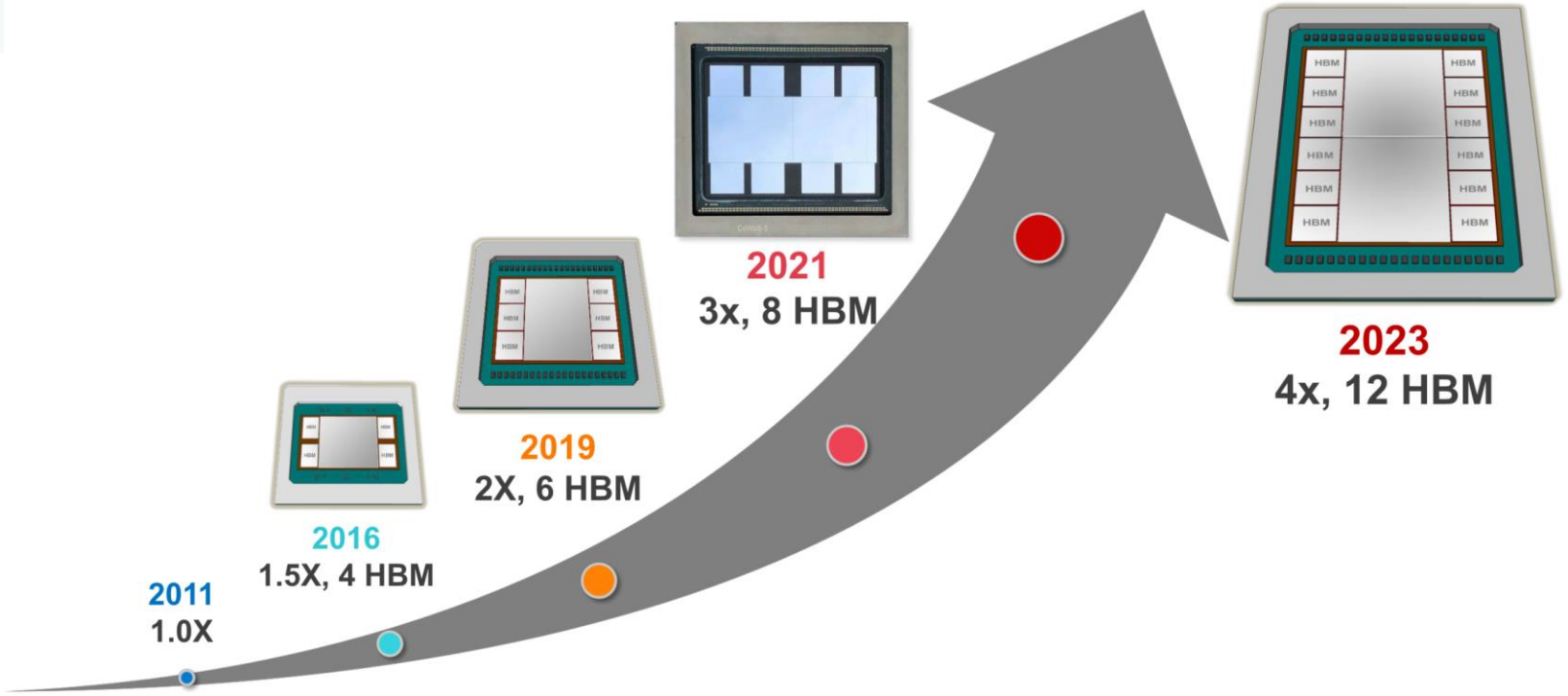
"Logic & Memory"

- **Silicon interposer offers the highest density connection**
  - Handle communication between HBM stack and logic (FPGA, GPU, or ASIC)
  - Communication between logic slices
  - Most mature, but expensive
- **Logic + HBM**
  - Early HBM = stacks of 4 DRAM + logic layer, now stack with 8 DRAMs common, future 12



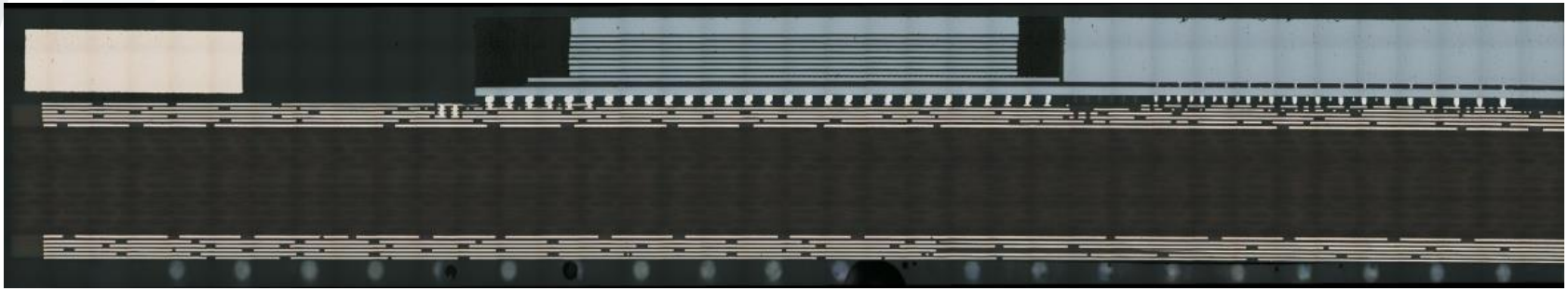
Source: Xilinx.

TSMC's Silicon Interposer Solution



- **Increasing size of interposer and multiple HBMs drives the size of the laminate substrate that forms the package**
  - Package body sizes of 65 mm x 65 mm, 75 mm x 75 mm, increasing to 100mm x 100mm body size
  - Applications include AI accelerators, FPGAs, network switch, GPUs

NVIDIA's GPU + HBM

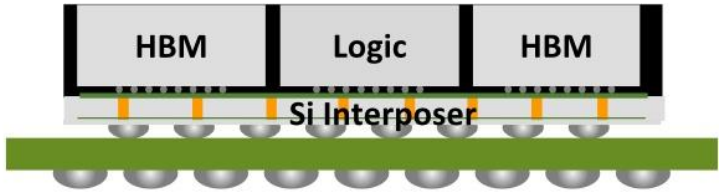


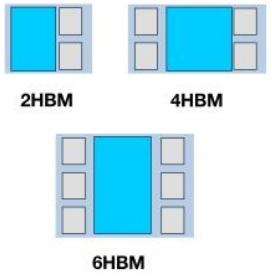
- **NVIDIA's GPU with 4 HBMs (8 high stack + logic layer) mounted on Si interposer**
  - HBM with wide bus (1,024 I/Os, ~4,000 bumps 55µm micro bump pitch)
- **NVIDIA's latest A100 uses GPU + 6 HBMs**
  - Continues to use CoWoS process

Source: NVIDIA.

Samsung I-Cube™ Silicon Interposer

- **Samsung's silicon interposer is part of its "cube" family of packages**
  - Demonstrated a 2,500mm^2 silicon interposer for up to 2 logic die plus 8 HBMs in an 85mm x 85mm package
  - Qualification for 2,800mm^2 silicon interposer is underway





Source: Samsung.

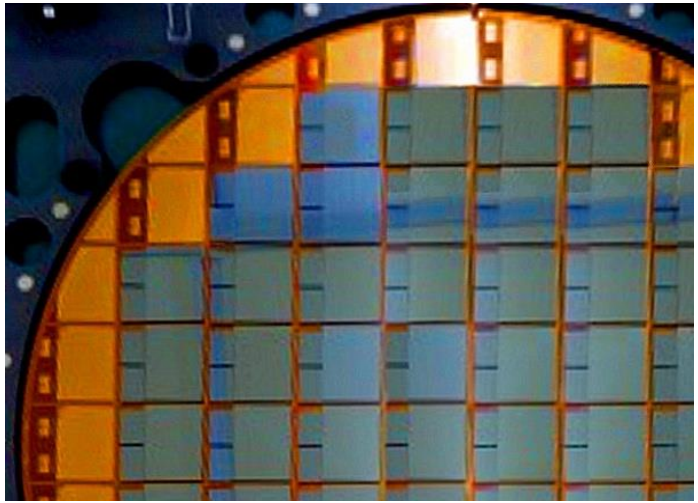
Why Fan-Out on Substrate Solutions?

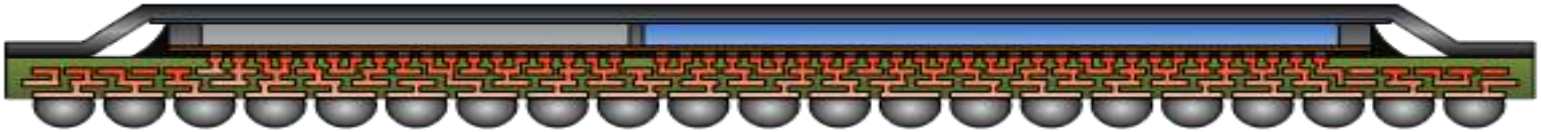
- **Fan-out on substrate applications in production for split die**
- **Many designs for HBM + logic under investigation with production expected in 2022-23**
- **Advantages of fan-out on substrate technology**
  - Lower cost than Si interposer
  - Potential for larger size packages than with Si interposer because there is less warpage
  - RDL with polyimide has CTE closer to that of the package substrate, resulting in lower residual stress at the package level

Source: TechSearch International, Inc., MediaTek.

ASE's FOCoS (Fan-Out Chip on Substrate)

- **Applications include network switch, GPU, AI**
  - In production since 2016 for split die for network switch
- **ASE chip first or chip last**
  - 2µm L/S
  - Up to 4 RDLs





FOCoS – Fan Out Chip on Substrate (FO FCBGA)

Source: ASE.

Amkor's Substrate-SWIFT®

- **Amkor's chip last FO version**
  - 2µm L/S
  - 4 RDLs qualified
  - 6 RDLs to be qualified by end of 2021

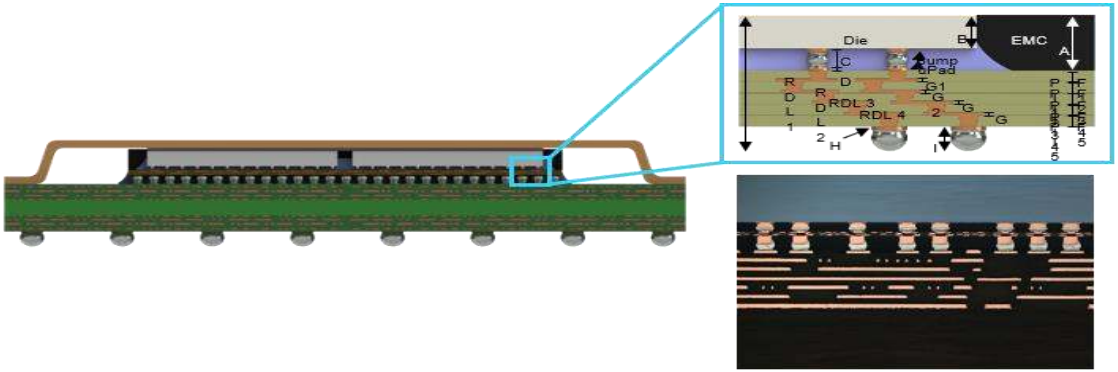
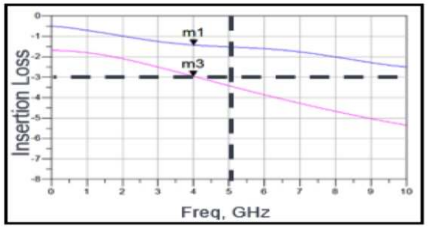
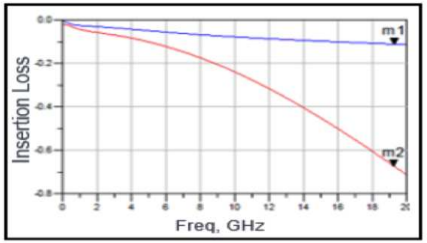


Image courtesy of Amkor Technology, Inc.



- Both 50Ω impedance
- M1 = SWIFT®
  - -1.42 dB loss @ 4 GHz
- M3 = 2.5D TSV
  - -2.95 dB loss @ 4 GHz

~4.5 mm trace length

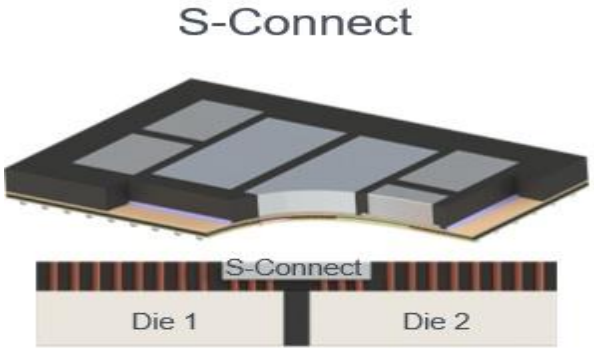


- S-SWIFT®
  - 4 RDL layers
  - Vias progression with dog-bones
- Silicon interposer
  - TSV (Through Silicon Via)

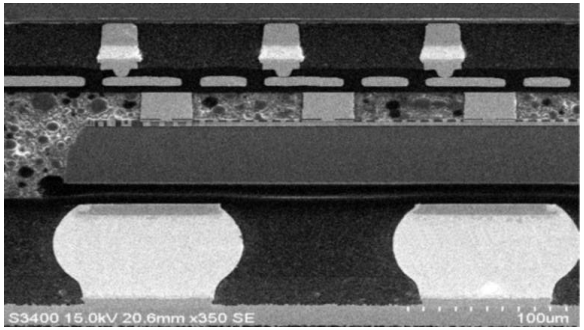
Images courtesy of Amkor Technology, Inc.

Why Embedded Bridge Solutions?

- **Advantages of embedded bridge technology**
  - Only use high-density connection where required
  - Potential for lower cost with reduced number of RDLs with FO RDL versions
  - Improved electrical performance
- **Many companies offer embedded bridge solutions**
  - Amkor (embedded in RDL, chip last)
  - ASE (embedded in RDL, chip last)
  - IBM (embedded in laminate substrate trench)
  - Intel (embedded in laminate substrate)
  - SPIL (embedded in RDL, chip last)
  - TSMC (embedded in RDL for FO, embedded with Si interposer-CoWoS)



Source: Amkor Technology

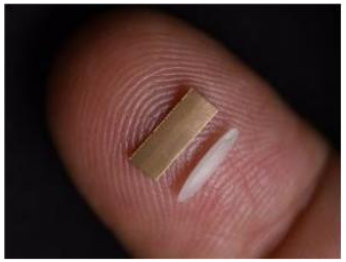


Source: ASE.

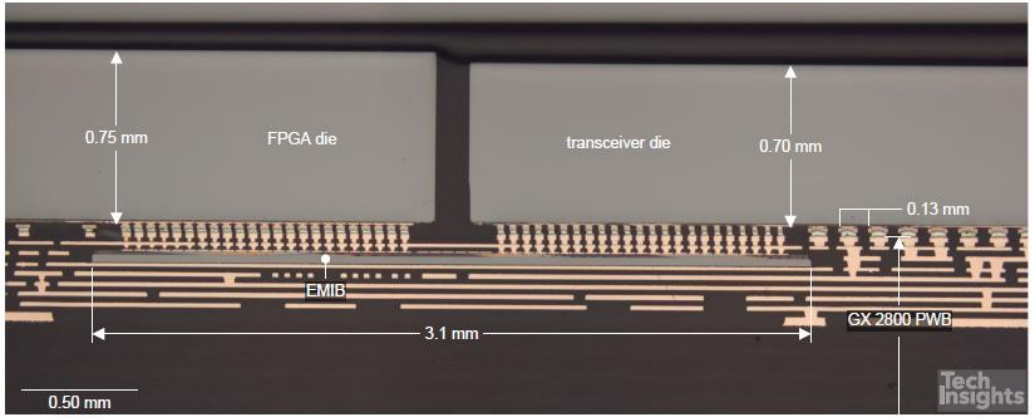
Intel's Embedded Multi-die Interconnect Bridge (EMIB)

Intel's Embedded Multi-die Interconnect Bridge (EMIB)

EMIB size compared to grain of Basmati rice



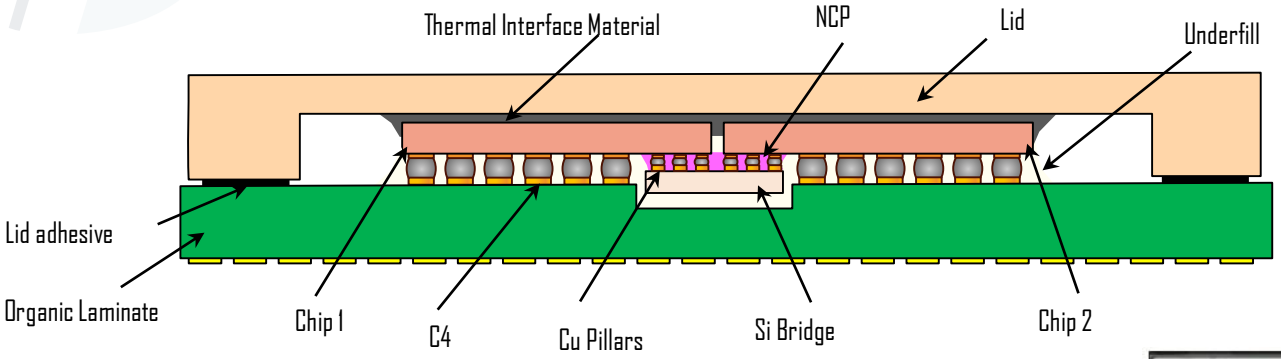
Source: Intel.

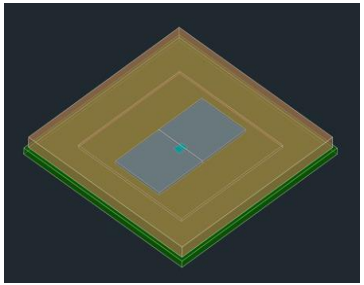


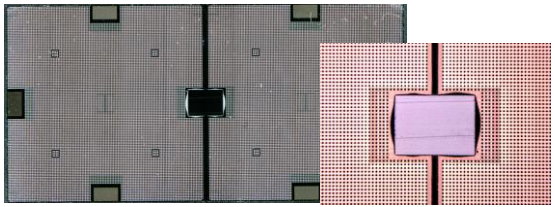
Source: TechInsights.

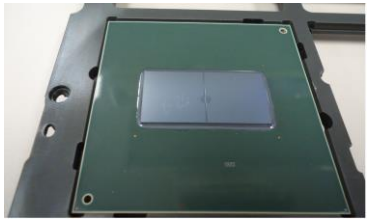
- **A small silicon bridge chip is embedded into the package (no TSVs) to provide in-package, high-density connection**
  - Micro bumps on chips, communication between chips through bridge die
  - Bridge die, designed and fabricated by Intel, embedded in the laminate substrate by substrate supplier
- **Provides a high-density localized interconnect between the FPGA and HBM, density ranges from 250 to 1,000 IO/mm/layer**
- **Applications include FPGA, discrete GPU, AI, server, and other data center HPC segments**

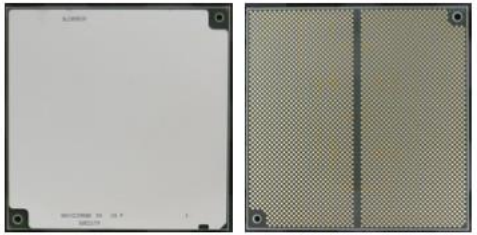
IBM Direct Bonded Heterogeneous Integration (DBHi) Si Bridge

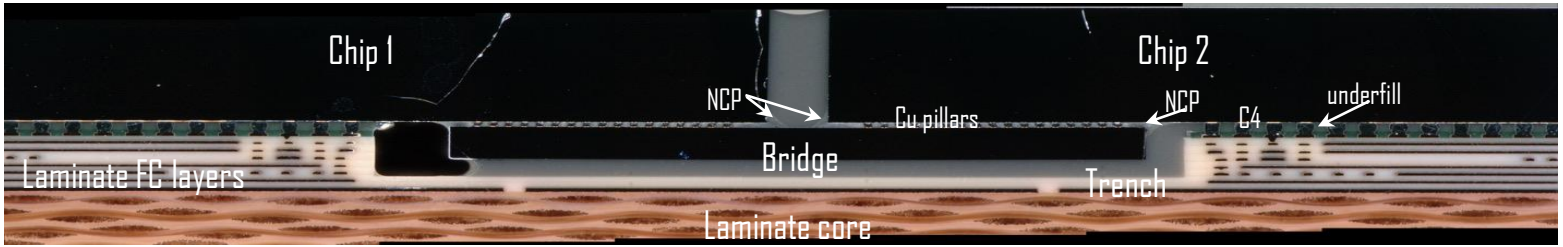






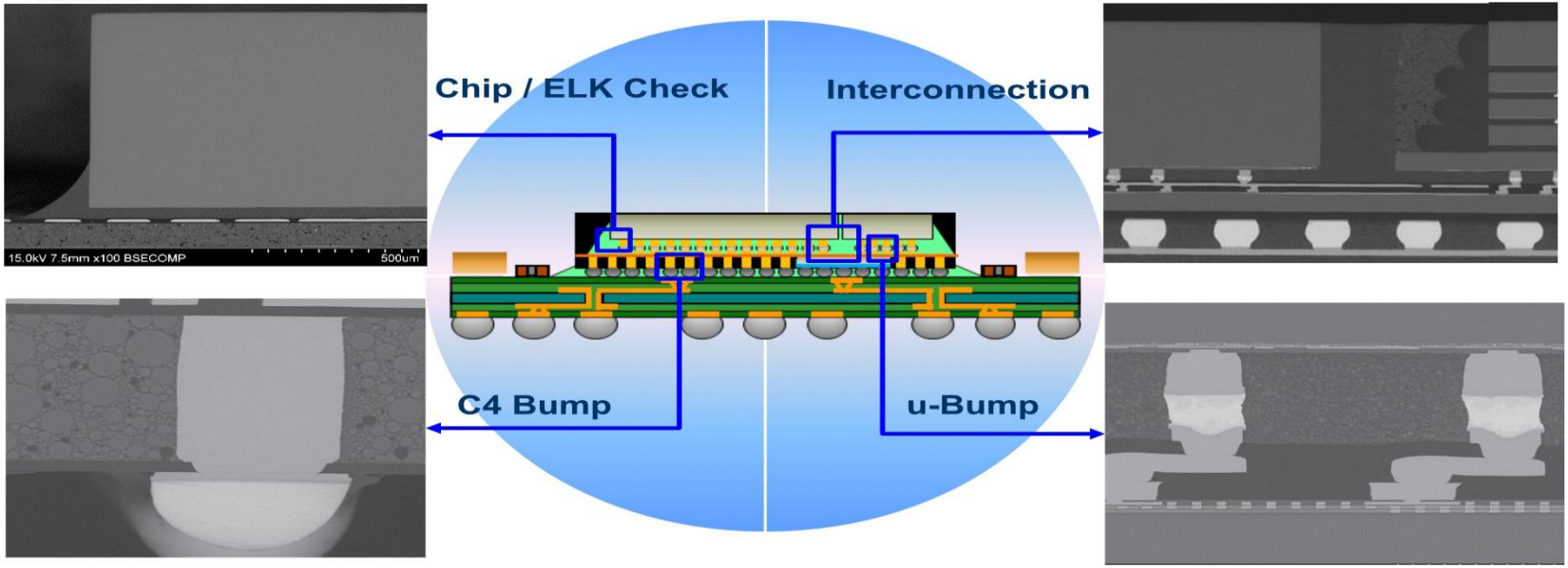






Source: IBM.

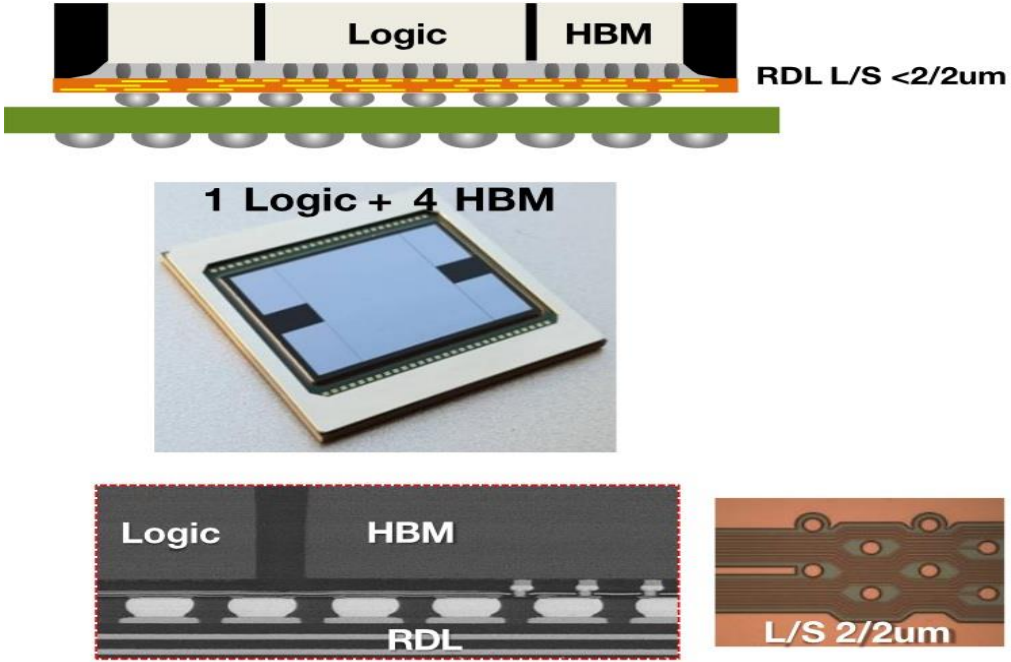
SPIL FO-EB Technology



Source: SPIL.

- **SPIL has qualified a fan-out embedded bridge (FO-EB) RDL test vehicle**
  - Microbumps used to connect each die to RDL
  - C4 bumps connect to next level
- **FO-EB targeted for CPU + HBM, GPU, and networking**

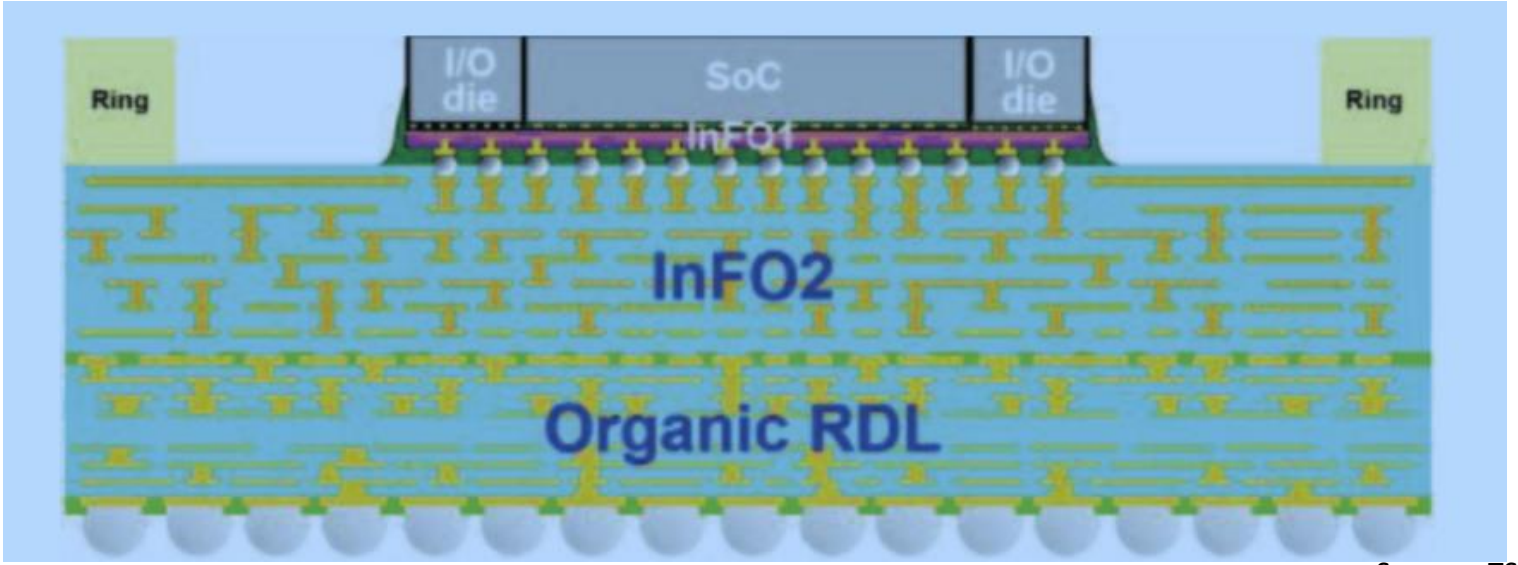
Samsung R-Cube™ RDL Interposer



Source: Samsung.

- **R-Cube is RDL interposer for logic + HBM**
  - Up to 4 layers have been fabricated
  - The 8mm x 12mm HBM2 has a 55µm minimum pitch
  - Test vehicle package is 55 mm x 55 mm

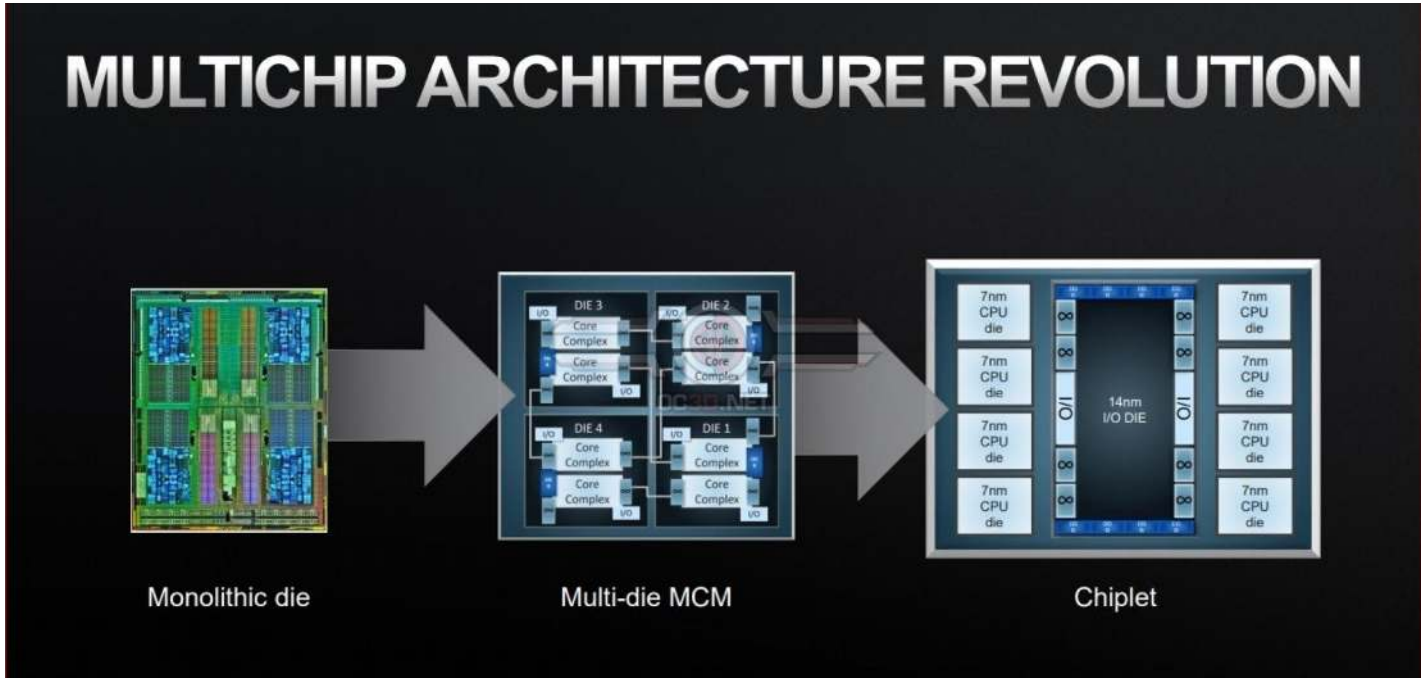
TSMC InFO System on Integrated Substrate (SoIS)



Source: TSMC.

- **High-density heterogeneous package using an RDL interposer with up to 6 interconnection layers**
  - Alternative to large, high-density substrates where yield and power consumption are challenges
- **Good electrical performance (large eye height, low jitter, and almost no layer-to-layer crosstalk), lower insertion loss**

New Era of Semiconductor Packaging

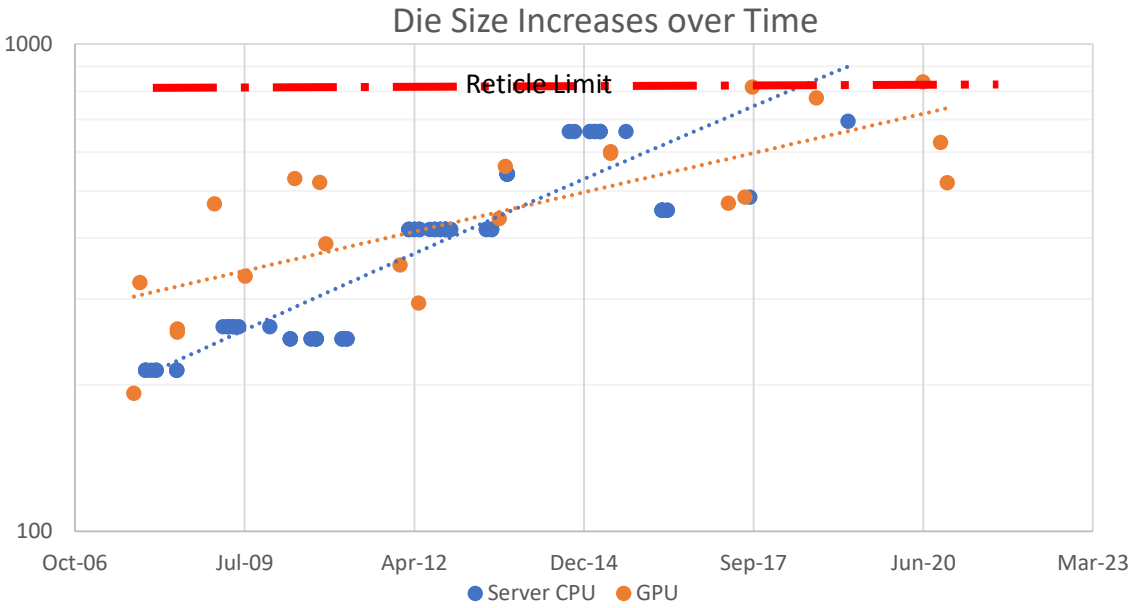


Source: Overclock3d.net.

- **Chiplets will be a key enabler for next 10-20 years**

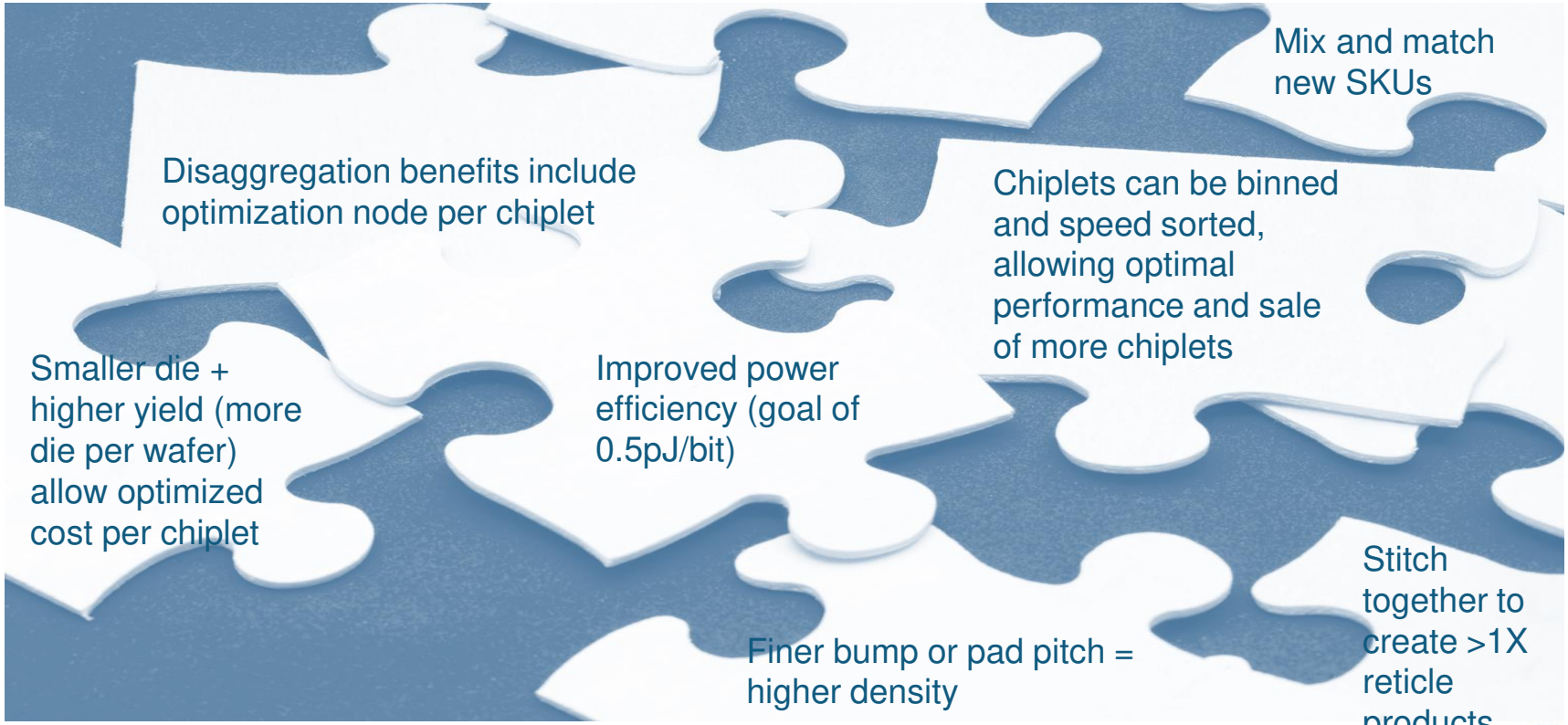
Die Size Growth: Major Driver for Adoption of Chiplets

- **Die sizes continued to increase over time for server CPU and GPU**
  - nVIDIA's 826 mm^2 die, fabricated at TSMC, is one of the largest in production today
  - Samsung reports die sizes of 750 mm^2
- **Performance requires more transistors, but industry needs a new, more economical approach**
  - Smart packaging, including heterogeneous integration and chiplets becomes the answer



Source: AMD internal analysis.

Many Drivers for Chiplet Adoption

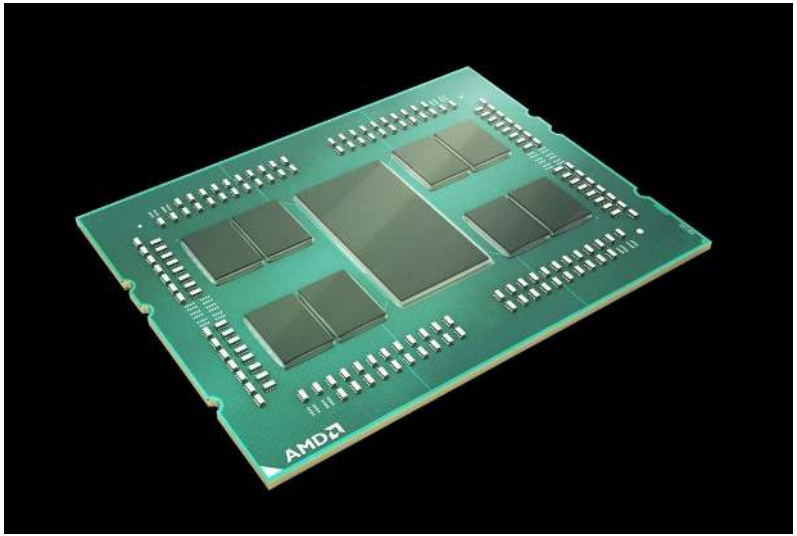


How Do We Define Chiplets?

- **A chiplet is an integrated circuit block specifically designed to work with other chiplets to form a larger more complex system that often makes use of reusable IP blocks**
  - A chiplet can be created by partitioning a die into functions that are more cost effectively fabricated (smaller die, higher yield, and less advanced nodes)
  - A chiplet is a hard IP block
  - Functions with other chiplets, so design must be co-optimized and silicon cannot be designed in isolation
  - Made possible by communication using chiplet interface (proprietary today)
- **Differs from SiP or traditional MCM in that it is a new design, not just a combination of different "off-the-shelf" chips**
- **Chiplet is not the package, it's the design philosophy**
  - Change from "silicon centric thinking" to "system-level planning" and "co-design of IC and package"
  - The industry has to think about chip design in a new way
  - Same impact as when the industry moved from a peripheral chip layout to area array!

AMD Multiple Chiplet Product Introductions

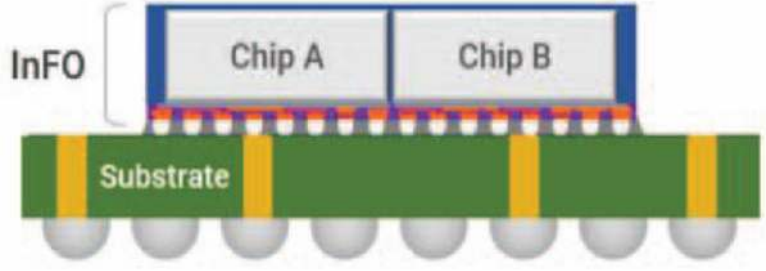
- **Multiple generations of desktop and server products using chiplets with organic substrate**
  - Split out analog functions from advanced 7nm logic
  - Chiplets can be binned and speed-sorted before assembly on the substrate
  - Better memory access
  - Minimize local latency
  - Power efficiency improvement
  - 1, 2, 4 or 8 CPU chiplets plus an I/O chiplet are attached to an organic interposer

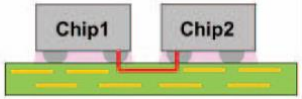
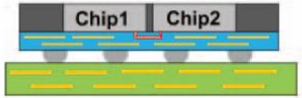


Source: Wired.com.

TSMC InFO_oS for Chiplets

- **Targets AI, network system, edge computing**
  - Alternative to the laminate MCM package for chiplets
- **FO process is a chip-first process**
  - Allows interconnect formation after wafer molding
  - Multiple layer RDL acts as a stress-buffer layer, so ELK stress is not a major concern
  - Process optimized to reduce warpage
- **Demonstration of 2.5x reticle of fan-out (51mm x 42mm) on a 110mm x 110mm substrate**
  - Test vehicle that integrates 10 chiplets (2 logic + 8 I/O die)
  - 5 RDLs (4 with 2/2µm, 1 with 5/5 µm)
  - D2D I/O pitch 36 µm



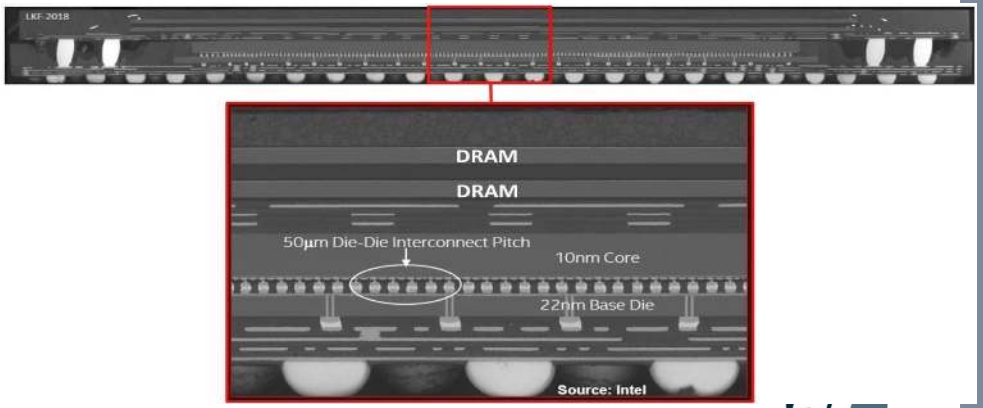
Technology	MCM	InFO_oS
Structure		
Min. Line W/S	15/15 um	2/2 um
Line counts/mm	34	250
BW/mm	1x	7.3x

Source: TSMC.

Intel Foveros Technology

- **Intel's Foveros technology die are stacked in 3D**
  - Base die, using less advanced node, can include power management features, voltage regulators, DC/DC converters
- **Benefits include**
  - Reduced voltage drop
  - Power efficiency
  - More immediate power delivery to the CPU cores
  - Elimination of passives on substrate
  - System-wide communication across multiple chiplets/dice vs. the limited die-to-die communication capability enabled by passive Si interposers
- **Used in the Samsung Galaxy Book S (Mobile PC)**
  - Longer lasting battery
  - No fan
  - Very thin package, allows thin product

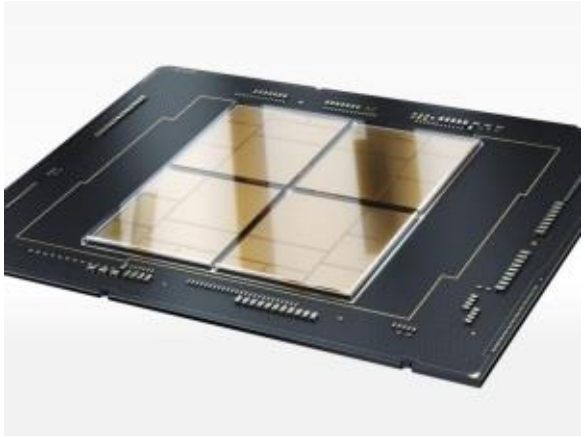
Intel's Lakefield CPU
- 10nm CPU
- 22nm Base die



Source: Intel.

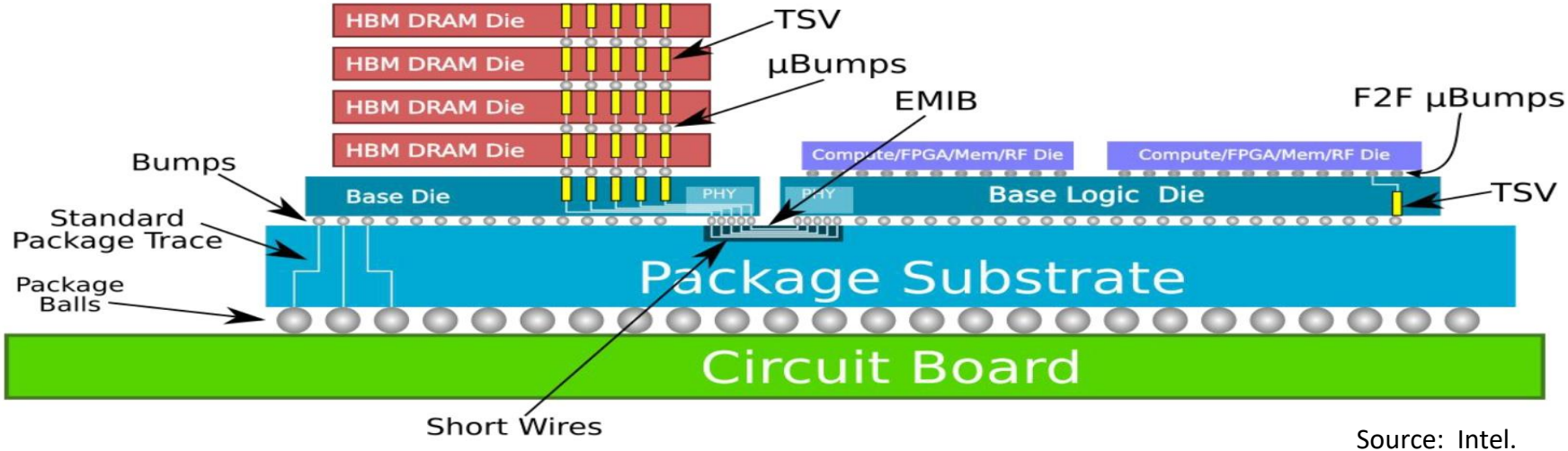
Intel's Data Center CPU with EMIB

- **Intel's Sapphire Rapids will be Intel's first CPU server for data centers using chiplets**
  - All 4 chiplet die will access shared cache and are connected using the modular die fabric (MDF)
  - Any core can talk to other cores on the 4 die and access the shared cache across all 4 quadrants as well as I/O across 4 quadrants



Source: Intel.

Intel Co-EMIB (EMIB + Foveros)



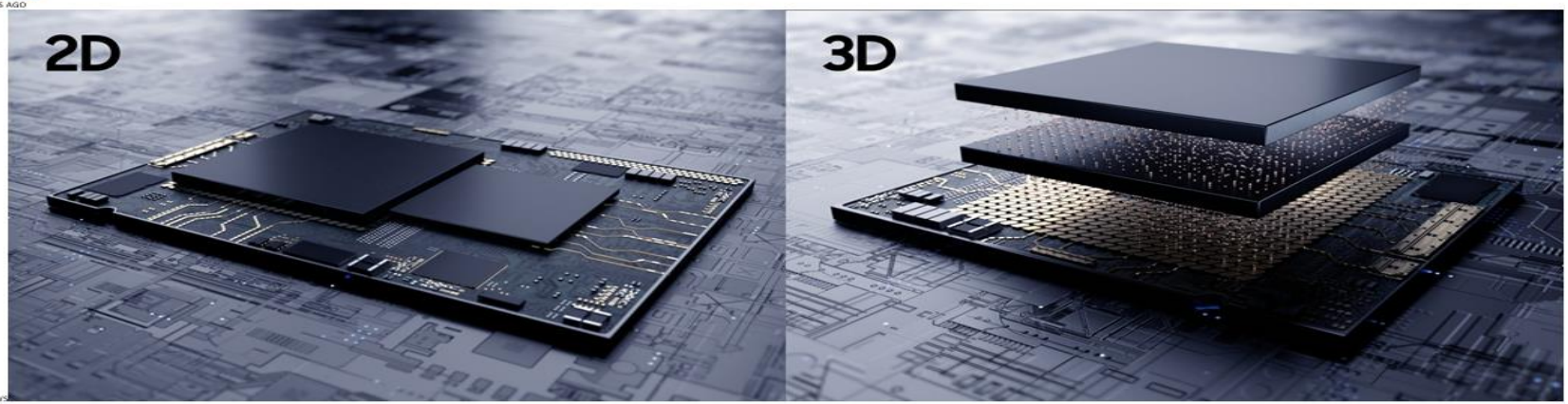
Source: Intel.

- **EMIB and Foveros can be combined to provide a high-density solution**
  - Connecting HBM and logic with silicon bridge
- **Intel's data center GPU with 2 large GPU 3D center tiles (Foveros), 8 HBM stacks and 2 additional chips each connected using EMIB**

Samsung 3D IC (X-Cube)

Samsung Announces Availability of its Silicon-Proven 3D IC Technology for High-Performance Applications

SHANNON DAVIS
3 HOURS AGO

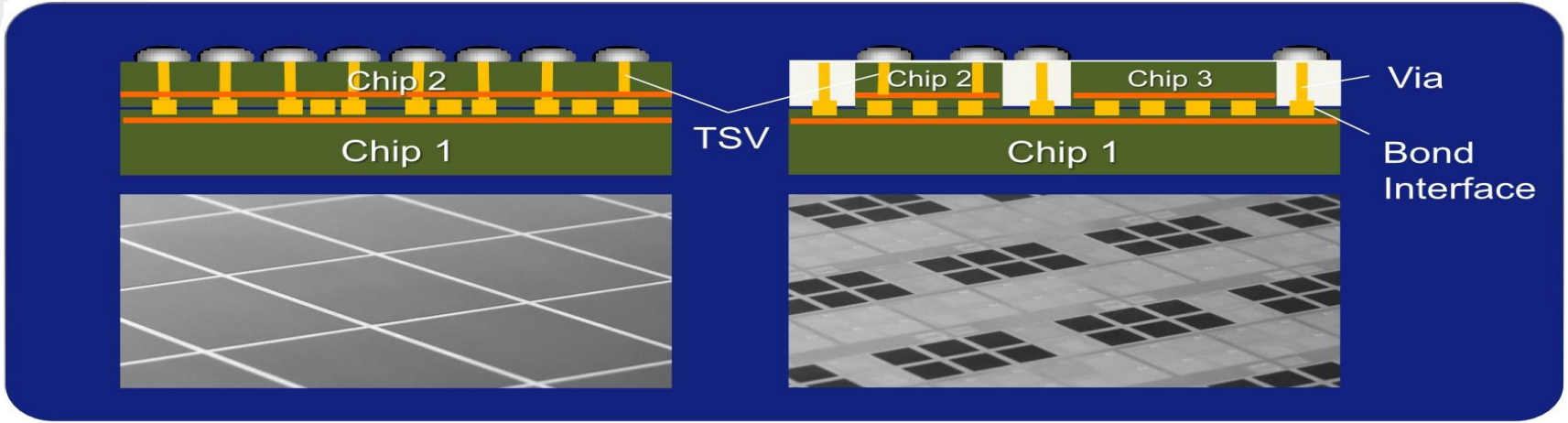


0 VIEWS

Source: Samsung.

- **Samsung introduction of logic and memory stack**
- **First version with µbump connections**

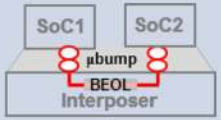
TSMC SoIC™ Technology



Source: TSMC.

- First commercial products with SoIC expected in 2021
- Commercial products with up to 10 chiplets expected in 2-3 years
- Advanced silicon nodes of 7nm or 5nm could be used
- SoIC could be placed next to HBM on RDL substrate or SoIC could be mounted next to HBM on CoWoS

SoIC™ Compared to 2.5D and 3D IC

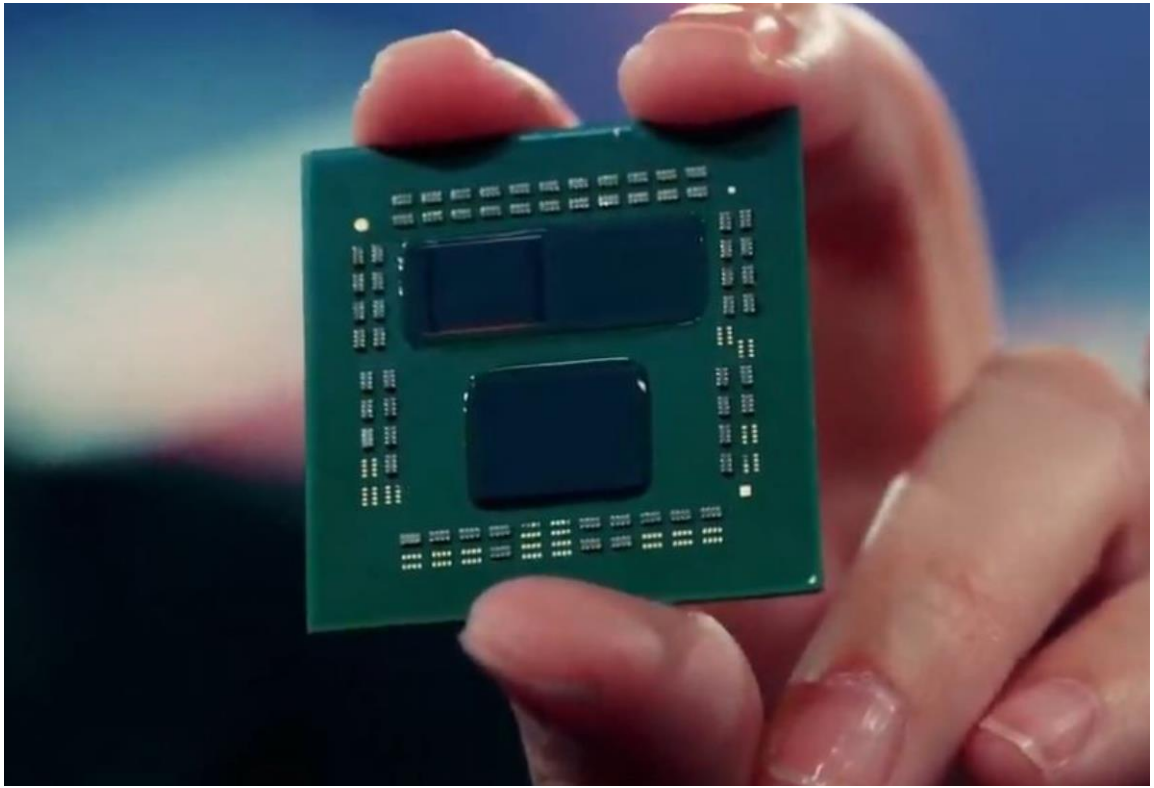
Technology	2.5D	3D-IC	SoIC
Structure cross-section			
Interconnect	µbump + BEOL	µbump	SoIC bond
Chip Distance	~100 µm	~30 µm	0
Bond-pad Pitch	36µm (1.0X)	36µm (1.0X)	9µm (0.25X)
Speed	0.01X	1.0X	11.9X
Bandwidth Density	0.01X	1.0X	191.0X
Power Efficiency (Energy/bit)	22.9X	1.0X	0.05X

Source: TSMC.

- **With SoIC there is virtually no distance between integrated chips, and a very small bond-pad pitch of 9 µm provides good scalability**
- **Using a bumpless bonding process is critical to improvements in performance, power, resistance, and capacitance (lower inductance and thermal resistance)**

AMD's 3D Chiplet

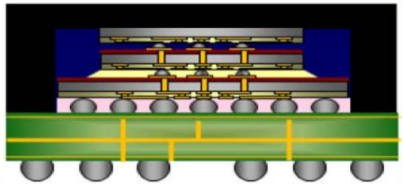
- **AMD's Prototype 5900X chip for gaming**
  - Gaming performance improvement
- **Same 7nm node as RYZEN, but performance gains using 3D chiplet copper-to-copper hybrid bond**



Source: AMD.

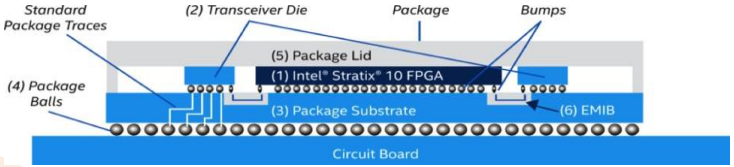
Many Package Options: Which One Do I Choose?

3D with µbumps



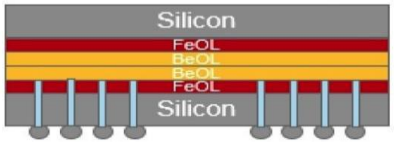
Source: Samsung.

Embedded Silicon Bridge



Source: Intel.

3D with hybrid bond



Source: TSMC.

Routing density requirements

Power efficiency and power delivery

Maturity of technology

Supply chain

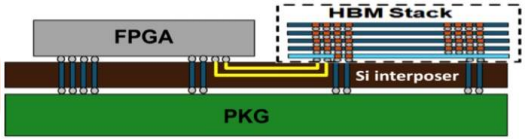
Thermal performance needs

Test considerations

Relative cost vs. alternatives

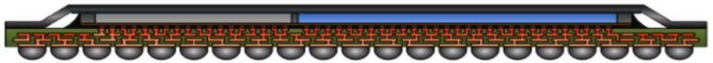
Product life and reliability requirements

Silicon Interposer

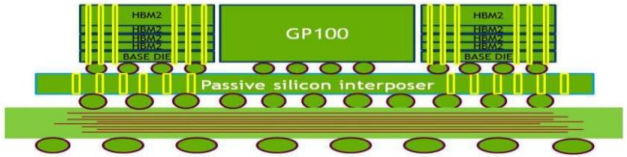


Source: Xilinx.

Fan-Out on Substrate



Source: ASE.



Source: Nvidia.

TechSearch INTERNATIONAL

Thank you!

TechSearch International, Inc.
4801 Spicewood Springs Road, Suite 150
Austin, Texas 78759 USA
+1.512.372.8887
tsi@techsearchinc.com

RELEVANT, ACCURATE, TIMELY