

封装摩尔时代的突破 -先进封装解芯片难题

证券分析师：唐仁杰 S0370524080002

研究员助理：谭景文

联系方式：tanjw1@jyzq.cn

行业评级：增持

摘要

- 什么驱动重心向先进封装领域倾斜？先进制程的成本呈现指数型增长，先进制程的“边际效益”下降（即随着关键尺寸微缩带来的边际成本下降）。一片2nm芯片的设计成本约7.25亿美元，是65nm芯片的25倍。此外，在CapEx方面，工厂建设和设备投入也观察到同样的现象，建造一座5纳米芯片制造厂所需的投资，是建造20纳米工厂的5倍。
- 从芯片设计及制造领域而言，芯粒及高端先进封装的组合可实现“混合制程”+缩短上市时间+可复用+良率改善。芯粒（Chiplet）基于需求考虑不同工艺，比如GPU需要较高性能选择3nm工艺，而I/O或模拟电路则可以使用成熟制程。再者，开发新产品可以复用此前IP，不需要整片IC设计，缩短研发周期及设计成本，并且能够实现独立验证。性能/瓦/美元（Perf/Watt/Dollar）综合来看，大芯片+3D堆叠更适合用于中小系统，而随着系统复杂度提升，“Small die with better yield”，即通过芯粒+3D堆叠的方式在大规模系统中性能/瓦/美元优势明显。
- 在单芯片原始计算性能方面，AI专用芯片（ASIC）弱于AI GPU。然而，即使是GPT-4这样的大语言模型也无法在单一芯片上运行。并且要达到与AI GPU相当的性能水平，ASIC需要构建比GPU集群更大规模的AI专用芯片集群。先进封装通过芯粒+异构实现更大面积拓展，这种可扩展性正是AI数据中心在控制成本的同时最大化性能的关键所在。在控制成本的同时，通过Chiplets+大中介层来突破尺寸限制，从而将AI加速器“做大做强”。光刻机的reticle（掩模版曝光视场）决定了单颗裸片在一次曝光里能做多大；超过这个面积就很难用传统单芯片（monolithic die）继续变大。
- 先进封装的技术演进核心是互连 I/O 数量与带宽密度持续上升：第一代以高密度电子互连为主，从 Si-Interposer、RDL-Interposer 到 EMIB/Co-EMIB（叠加 TSV），再到中介层与桥接器等形态，在微凸块/微铜柱基础上实现从存内堆叠到存算堆叠与算算堆叠的更高集成。随着间距继续缩小，混合键合（Hybrid Bond）成为提升互连密度与能效的关键。进入第二代，封装不再只通过电连接，而是把光互连引入封装体系（Fiber Optic + 电子互连），面向“未来封装=小芯片（Chiplets）+ 异构集成 + 光学 I/O”的方向演进，以支撑 AI 时代更高的互连 I/O 需求并缓解带宽与功耗瓶颈。
- 硅桥封装技术是一种2.5D解决方案，用于替代硅中介层技术。其核心是将一个或多个硅桥集成在特定的封装基板（可由不同材料制成）或模塑中介层中，以确保两个或多个芯片之间的互连。硅桥可以封装在基板上或内以及模塑体内，各公司工艺略有差异。嵌入式把硅桥放在基板腔体里再做布线，硅桥与基板过渡更像同一平面系统。而将硅桥封装在模塑体内，其布线密度要高于基板的方案。
- 制约2.5D互连密度的主要由三个因素构成：焊料桥接（Solder Bridging）风险、金属间化合物（Intermetallic Compounds, IMC）、底部填充（Underfill）工艺的挑战。通过直接键合（Direct Bonding）和混合键合（Hybrid Bonding）实现“去焊料化”从而实现互连密度提升是3D封装的关键。混合键合技术通过在原子尺度上实现电介质与金属的直接连接，消除了焊料层，从而将互连间距从微米级（20 μm）推升至<10 μm。
- W2W、D2W、Co-D2W：W2W是指将两整片晶圆（通常均为300mm规格）进行整面对准、键合，随后进行减薄、TSV露头及切割的工艺流程。这是目前混合键合技术中最成熟、应用最广泛的形态；D2W是将经过测试、切割后的独立裸片（KGD），逐个拾取并以高精度键合到目标晶圆（Target Wafer）的特定位置上，是实现高性能异构集成的方案。Co-D2W 是一种折衷方案，旨在结合W2W的高效率和D2W的KGD优势，将KGD集体排放后同时键合，提高D2W吞吐。
- 先进封装市场方面，2024年中国先进封装市场约967亿元，占全球市场规模的30.95%，随着AI对高性能算力芯片的需求，预计2029年中国半导体先进封装测试市场将达到1888亿元，2024-2029年年复合增速达14.30%，2029年预计中国先进封装市场将占全球市场规模36%。从单位封装成本来看，不含基板价值量中由于使用Si中介层、模塑中介层嵌入硅桥技术良率、工艺复杂度相对较高，单位价值量较高。存储应用中，HBM整体高于CBA DRAM、3D NAND闪存。
- 相关公司：1、设备厂商：拓荆科技（688072.SH）、中微公司（688012.SH）、盛美上海（688082.SH）、光力科技（300480.SZ）、北方华创（002371.SZ）、中科飞测（688361.SH）等；2、材料：鼎龙股份（300054.SZ）、安集科技（688019.SH）、飞凯材料（300398.SZ）等；3、OSAT：盛合晶微（未上市）、长电科技（600584.SH）、深科技（000021.SZ）等
- 风险提示：1、技术风险：当前2.5D封装及3D封装种类较多，分歧较大，可能导致相关公司业绩受到影响；2、供应链风险：当前封装设备及零部件国产化率不足，导致设备供应不及预期；3、AI需求不及预期：国内数据中心算力需求不及预期，导致产业链需求不及预期。



目录

● 一、先进封装的摩尔定律：
解成本

● 二、先进封装的摩尔定律：
解拓展

● 三、先进封装的摩尔定律：
解互连

● 四、先进封装的摩尔定律：
解边缘AI

● 五、先进封装技术演进-
2.5D

● 六、先进封装技术演进
-3D

● 七、先进封装市场-
2.5D、3D工艺价值量高

● 八、相关公司

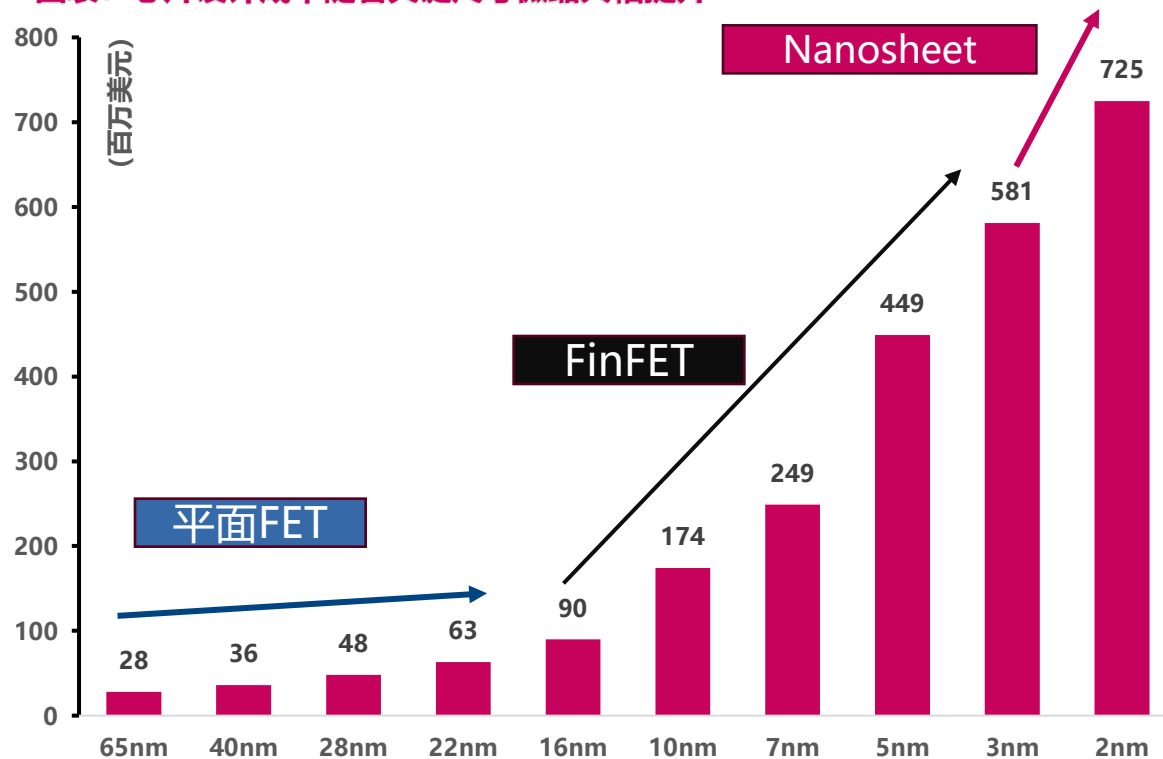
风险提示：

- 技术风险：当前2.5D封装及3D封装种类较多，分歧较大，可能导致相关公司业绩受到影响
- 供应链风险：当前封装设备及零部件国产化率不足，导致设备供应不及预期
- AI需求不及预期：国内数据中心算力需求不及预期，导致产业链需求不及预期

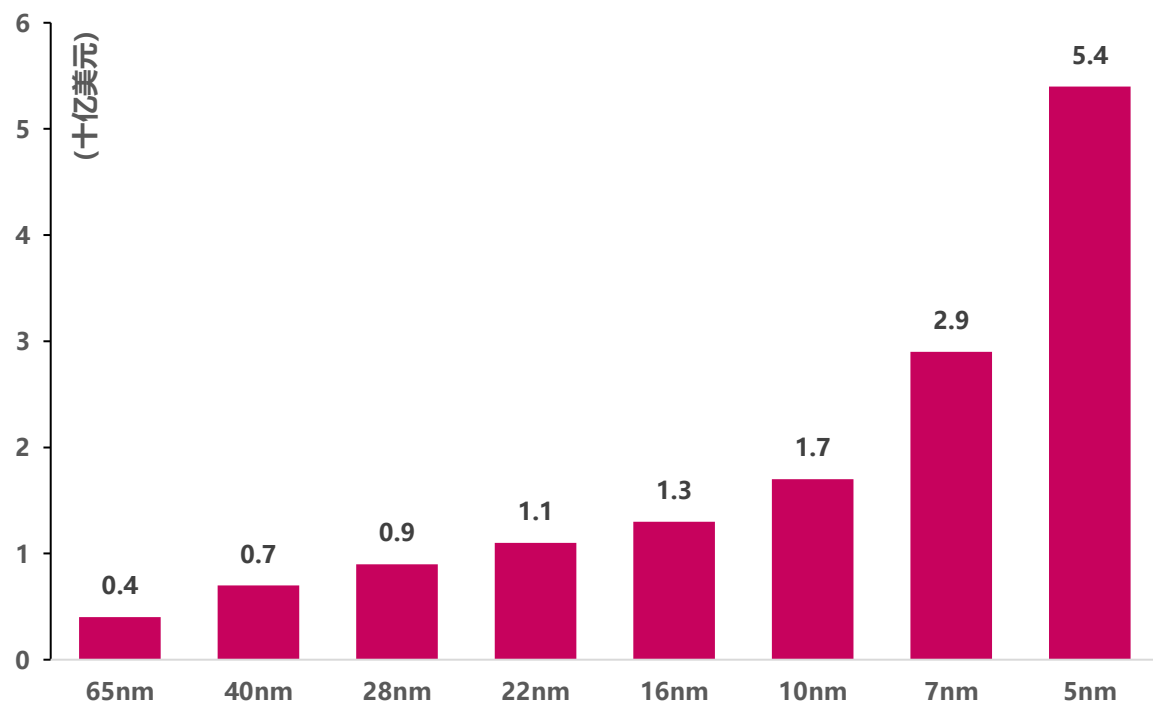
先进封装的摩尔定律：解成本

- 什么驱动重心向先进封装领域倾斜？先进制程的成本呈现指数型增长，先进制程的“边际效益”下降（即随着关键尺寸微缩带来的边际成本下降）。
- 在设计方面，一片2nm芯片的设计成本约7.25亿美元，是65nm芯片的25倍。尤其是当FET工艺由平面转向FinFET及Nanosheet后，随着制程节点不断缩小，由于量子效应、微小的结构变异以及测试、验证和IP认证等因素带来的额外困难，开发工作变得愈加复杂，这进一步加剧了开发流程的复杂性和进度延迟。
- 此外，在CapEx方面，工厂建设和设备投入也观察到同样的现象，建造一座5纳米芯片制造厂所需的投资，是建造20纳米工厂的5倍。

➤ 图表：芯片设计成本随着关键尺寸微缩大幅提升



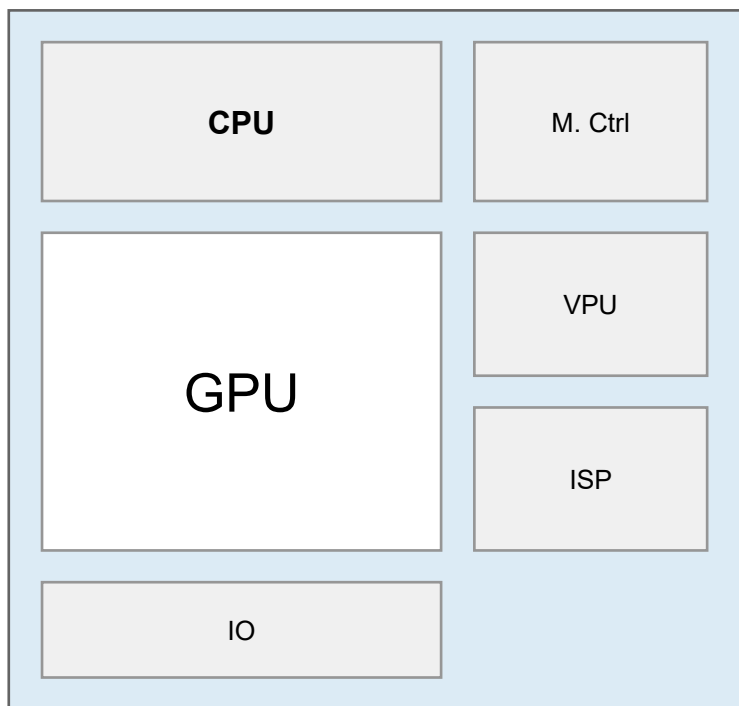
➤ 图表：同样，Fab及设备投入也呈现指数型增长



先进封装的摩尔定律：解成本

➢ 从芯片设计及制造领域而言，芯粒及高端先进封装的组合可实现“混合制程”+缩短上市时间+可复用+良率改善。芯粒（Chiplet）基于需求考虑不同工艺，比如CPU需要较高性能选择3nm工艺，而I/O或模拟电路则可以使用成熟制程。再者，开发新产品可以复用此前IP，不需要整片IC设计，缩短研发周期及设计成本，并且能够实现独立验证

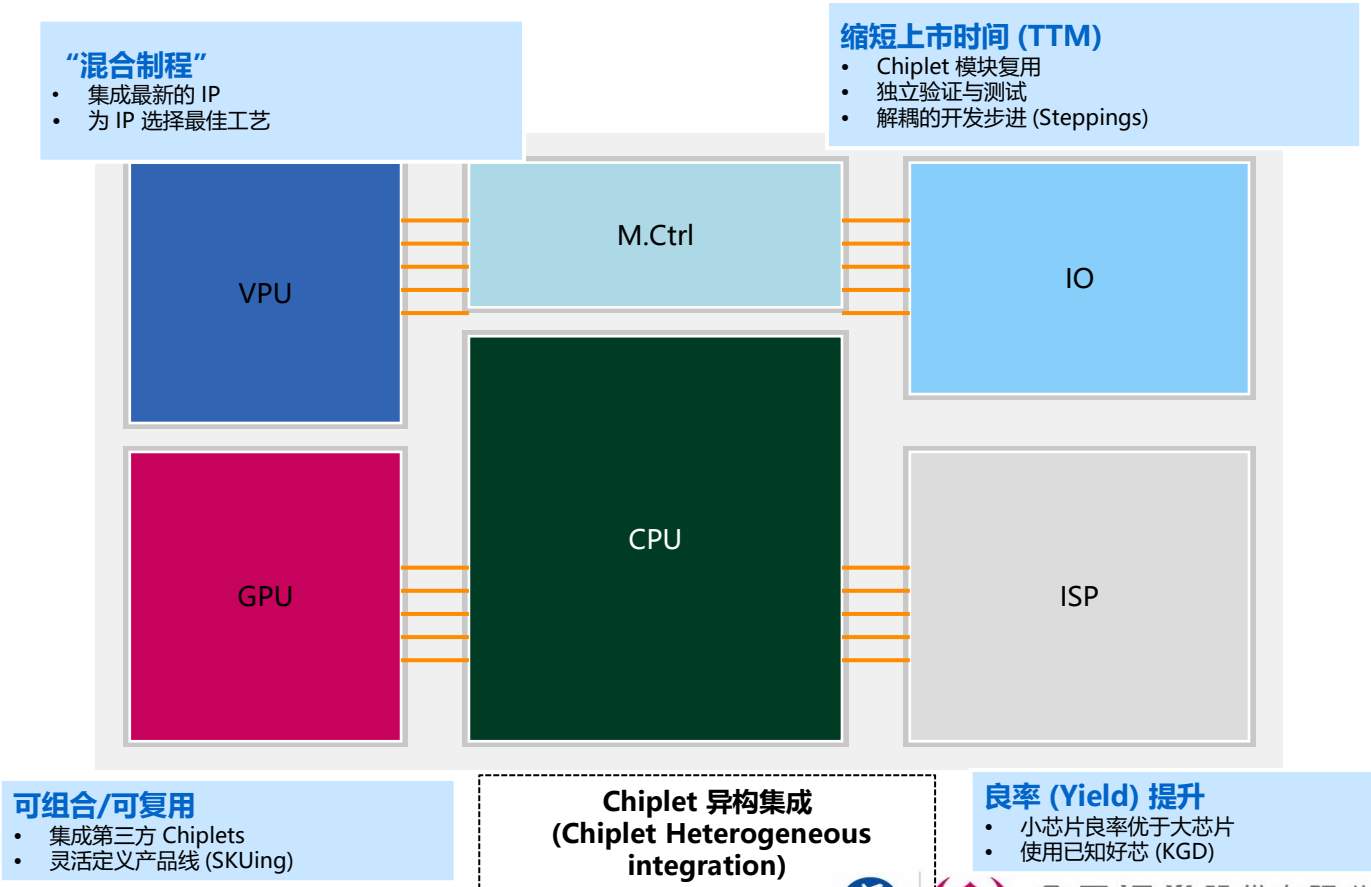
➢ 图表：单片集成



单片集成 (Monolithic Die)

- 将所有子系统集成在同一裸片、
- 使用同一工艺节点制造。

➢ 图表：芯粒异构集成，不同功能芯片解耦制程，降低成本、良率提升且缩短上市周期



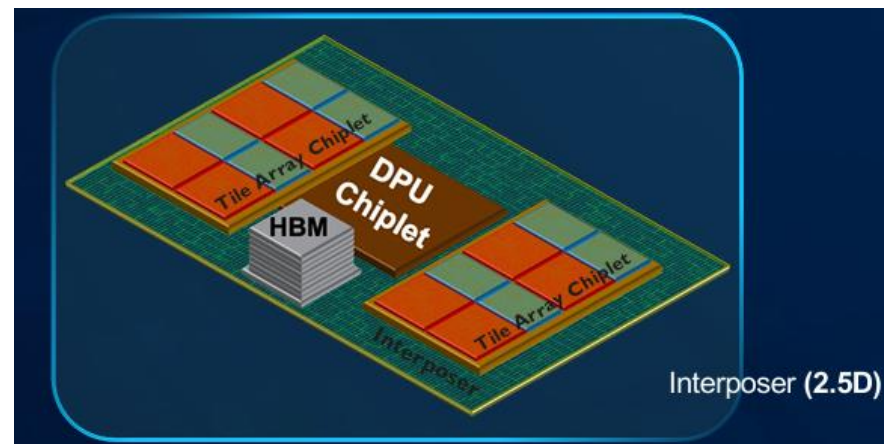
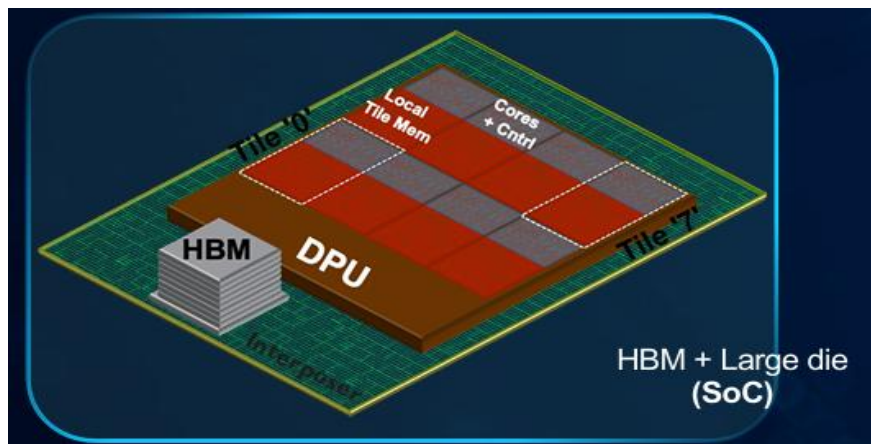
先进封装的摩尔定律：解成本

- 考虑四种不同架构设计：SoC（单片）、2.5D（中介层封装）、L3D（大芯片3D堆叠）与 S3D（系统级3D堆叠）。从技术上而言，SoC是将CPU、GPU、IO等所有模块采用同一工艺制造，而Chiplet异构集成通过将大芯片拆解成独立的模块（Tiles），再通过2.5D封装或者3D堆叠的方式进行封装。

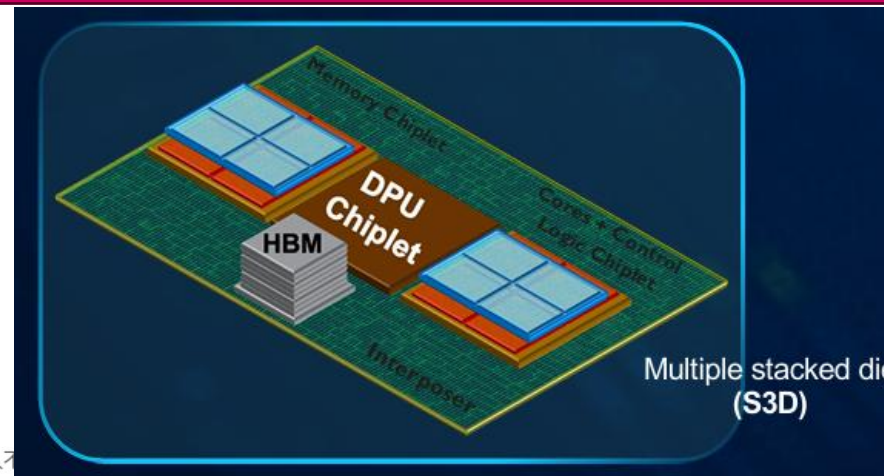
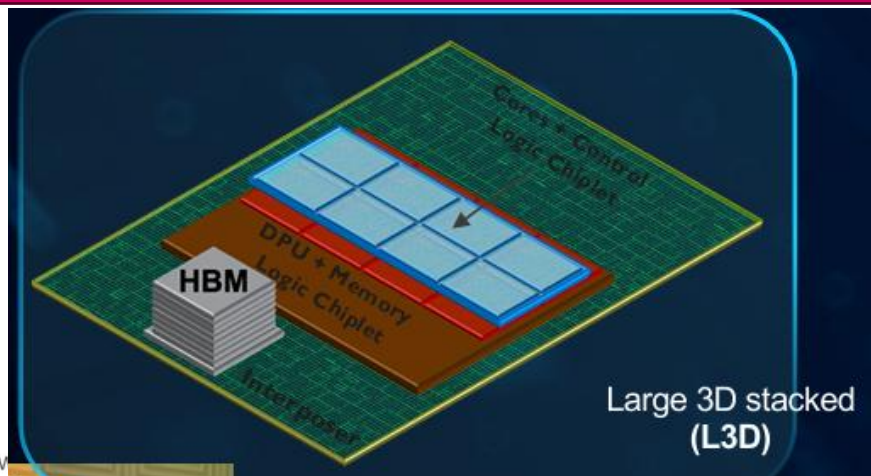
极少芯粒

更多芯粒

不进行3D堆叠



进行3D堆叠



www

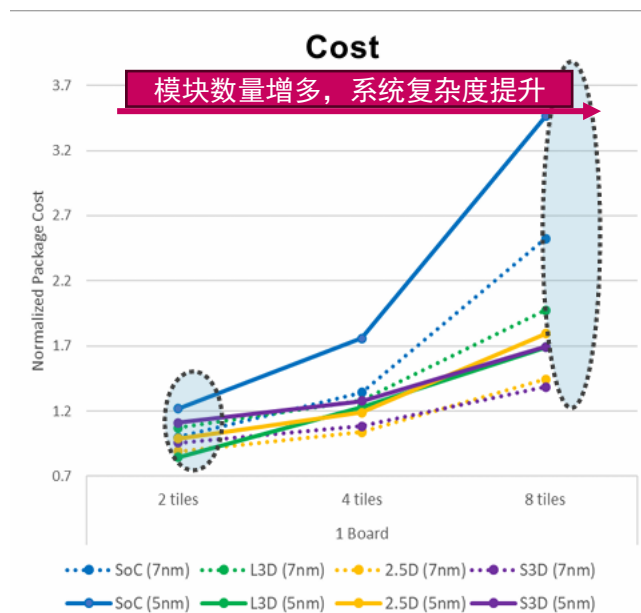
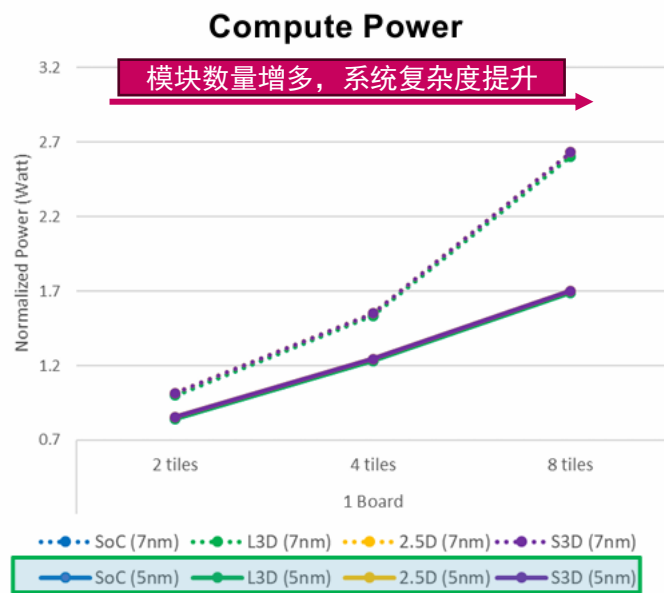
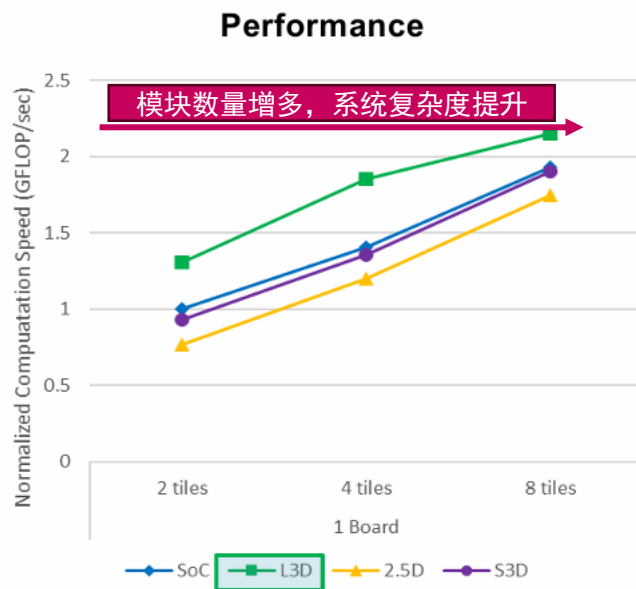
有, 未经许可, 单位或个人

有限公司
ES CO.,LTD.

先进封装的摩尔定律：解成本

- 单一从性能上而言，3D 堆叠技术利用 TSV（硅通孔）实现了最短的垂直互连距离，消除了 Chiplet 之间的横向通信延迟（Inter-chiplet latency）。相比之下，SoC及2.5D、S3D 受限于物理布线长度，性能略逊一筹。但可以看出，随着系统复杂度的提升，多芯粒及3D堆叠的S3D性能与SoC差异较小，不过成本上SoC要比S3D高的多。
- 从功耗上而言，制程的优势较为显著。架构带来的差异（SoC vs 3D）远小于工艺升级带来的红利。虽然 3D 堆叠本身会引入微小的散热和供电挑战，但先进工艺的低漏电特性弥补。
- 芯粒+3D堆叠的成本优势在系统复杂度提升下愈发明显。当系统规模达到 8 Tiles 时，SoC 的成本呈指数级暴涨。这是因为超大芯片的良率大幅下行，制造一块好芯片需要废弃无数块晶圆。而S3D则通过小芯片+3D堆叠从而实现近似性能下的成本优势。

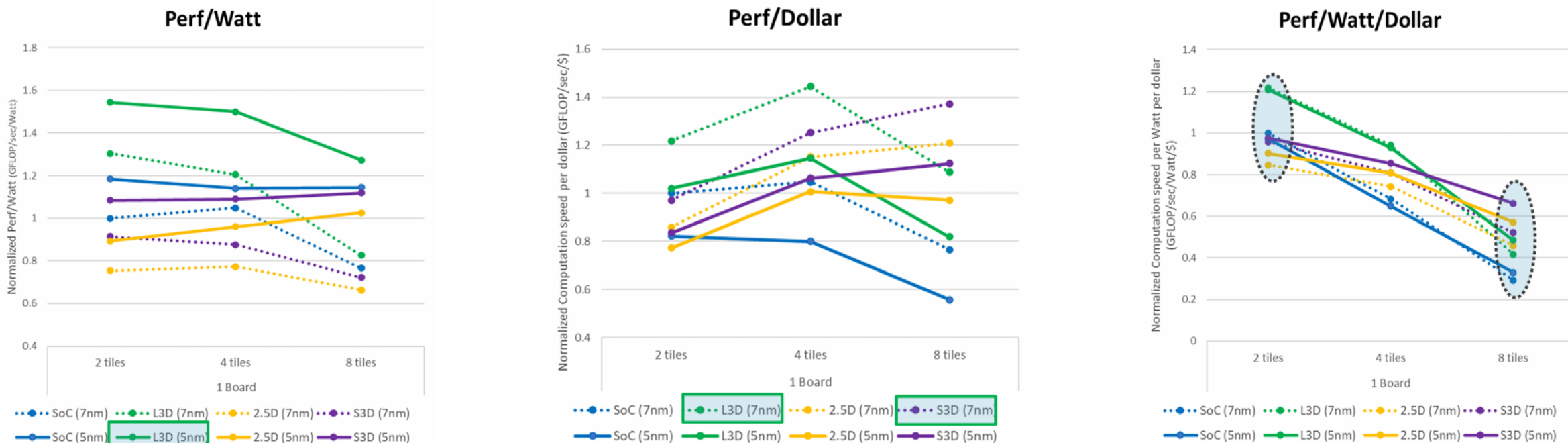
➤ 图表：不同封装/制程下的性能、功耗、成本对比



先进封装的摩尔定律：解成本

- 同时考虑性能、成本及功耗，L3D在小系统层面具有明显的性能优势，每瓦性能（Perf/Watt）领先，而在性价比角度，性能/美元（Perf/Dollar）指标中S3D（7nm）在大系统中表现最好，成本敏感的大型系统，通过成熟工艺+多重堆叠是更好的选择。
- 性能/瓦/美元（Perf/Watt/Dollar）综合来看，大芯片+3D堆叠更适合用于中小系统，而随着系统复杂度提升，“Small die with better yield”，即通过芯粒+3D堆叠的方式在大规模系统中性能/瓦/美元优势明显。

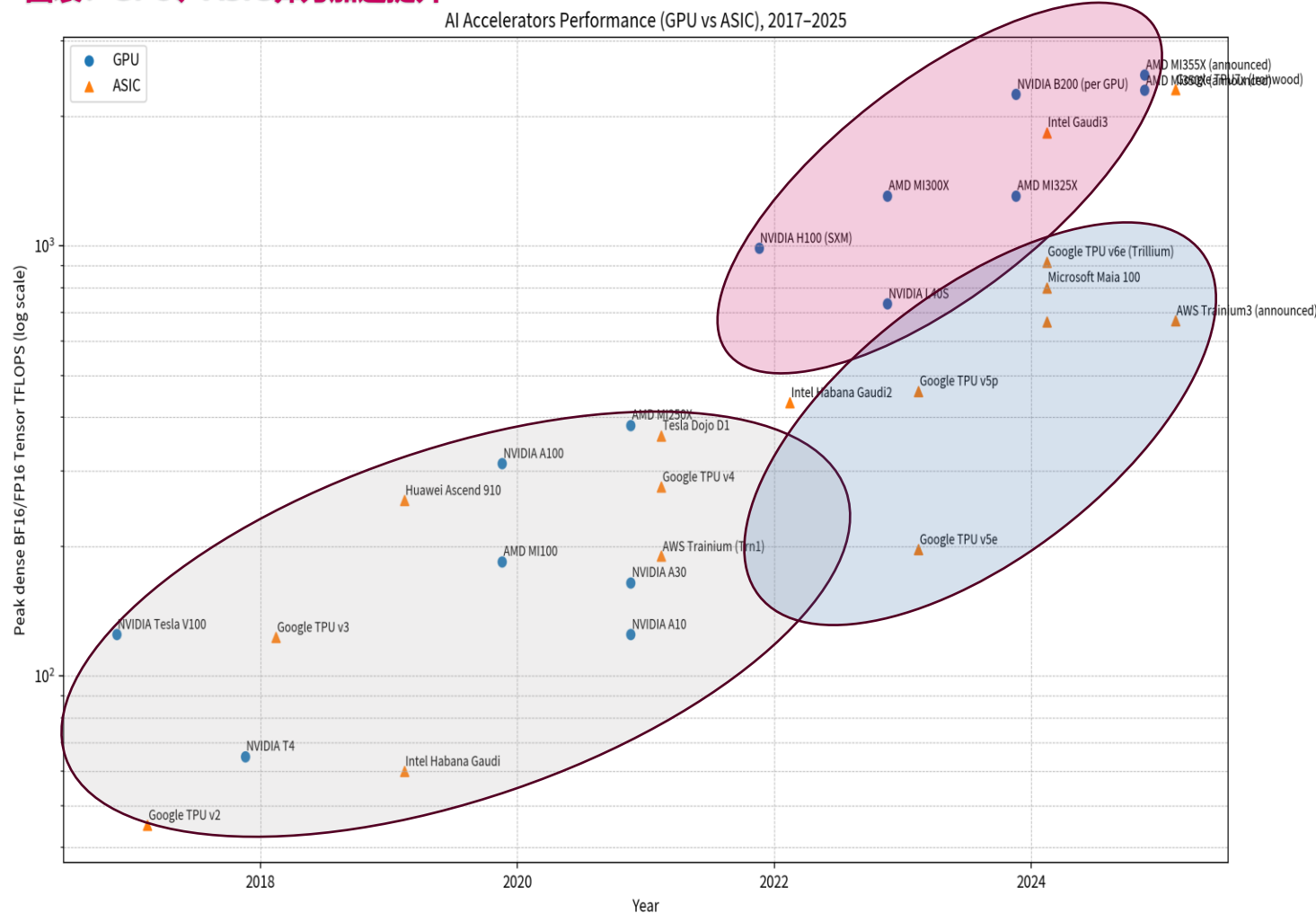
图表：中小系统中，大芯片+3D堆叠占优，随着系统复杂度提升，芯粒+3D堆叠占优



先进封装的摩尔定律：解拓展

- **AI训练及推理对性能的追求加速AI加速器的发展速率**：据统计，在2017年至2022年期间，GPU及ASIC的单芯片计算性能以年均47%的速度增长；而2022年底ChatGPT的发布加速了这一趋势，目前年增长率已提升至约84%。
- **Scaling Law背景下，GPGPU的通用计算能力是AI训练的首选加速器**，训练芯片需要同时具备巨大的算力和内存带宽。2.5D封装集成HBM成为业界标配方案，通过硅中介层将HBM存储器紧贴GPU/加速器，实现数TB/s级内存带宽以缓解内存墙。
- 此外，在单芯片原始计算性能方面，AI专用芯片（ASIC）弱于AI GPU。然而，即使是GPT-4这样的大语言模型也无法在单一芯片上运行。并且要达到与AI GPU相当的性能水平，**ASIC需要构建比GPU集群更大规模的AI专用芯片集群**。先进封装通过**芯粒+异构实现更大面积拓展**，这种可扩展性正是AI数据中心在控制成本的同时最大化性能的关键所在。

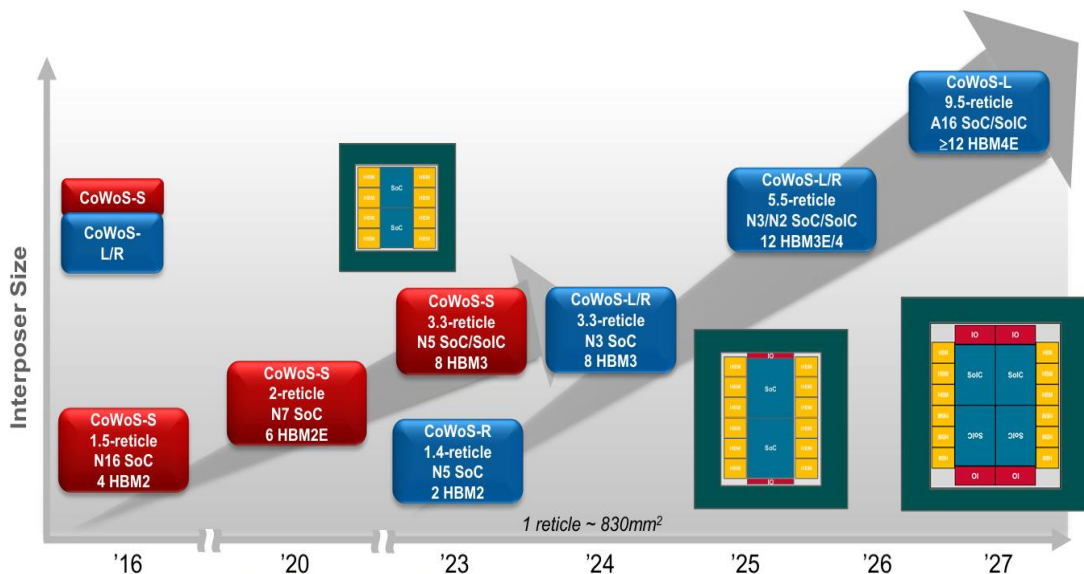
➤ 图表：GPU、ASIC算力加速提升



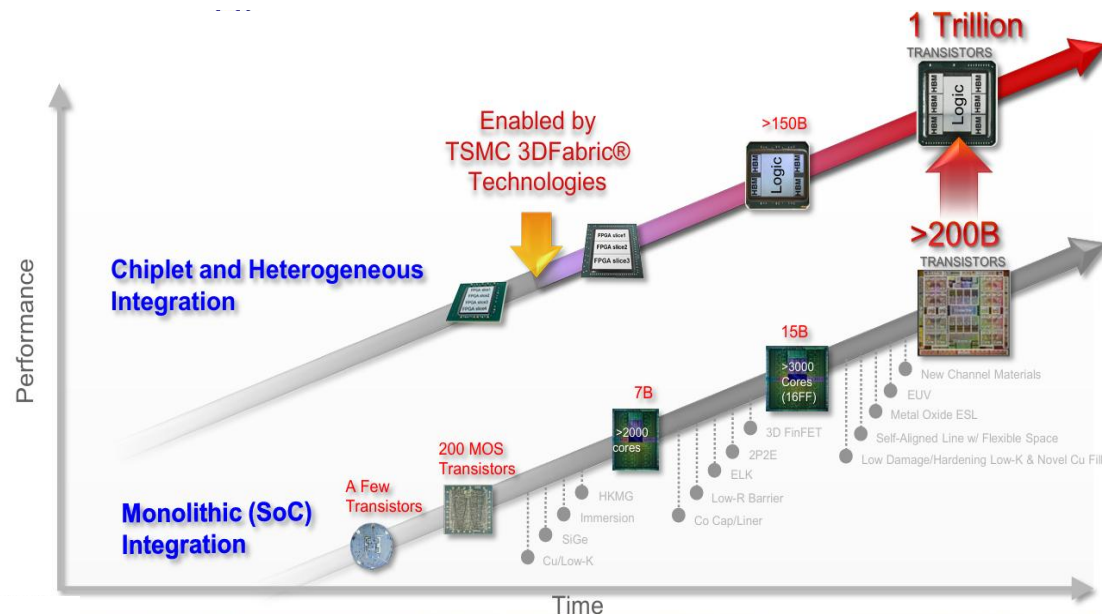
先进封装的摩尔定律：解拓展

- 在控制成本的同时，通过Chiplets+大中介层来突破尺寸限制，从而将AI加速器“做大做强”。光刻机的 reticle（掩模版曝光视场）决定了单颗裸片在一次曝光里能做多大；超过这个面积就很难用传统单芯片（monolithic die）继续变大。AI芯片恰恰既要更大的算力阵列，又要更高的存储带宽与容量，所以继续靠把一颗芯片做得越来越大，很快就会被 reticle 上限、良率与成本限制。
- 以台积电的CoWoS-L/R为例，将计算逻辑芯片拆成若干个仍在reticle允许范围内的SoC/Chiplet，再将其放置在有机中介层。中介层可以通过多次曝光拼接/扩展成更大面积（如2-reticle、3.3-reticle等），整个封装面积跨越了reticle限制，而计算芯粒（Chiplets）本身保持在可制造、可控成本与良率的尺寸区间内。随着中介层承载面积增大，可承载芯粒数量也将得到扩张。

➤ 图表：先进封装突破单芯片尺寸上限










➤ 图表：通过芯粒、异质集成实现更强性能



先进封装的摩尔定律：解拓展

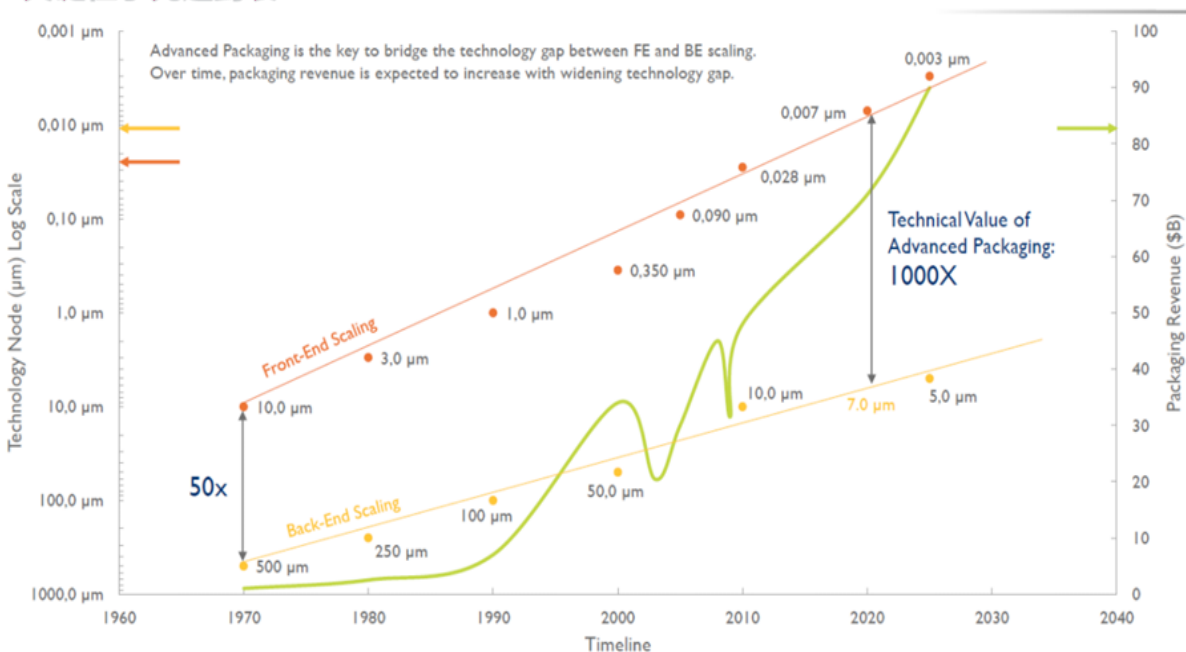
➤ 当前主流GPU、ASIC均采用2.5D封装，不过在中介层上略有差异。未来由于硅中介层的尺寸限制或转向RDL (R) 中介层及嵌入硅桥 (L)

公司	型号	处理器类型	使用的HBM类型	内存总容量 (GB)	单堆栈容量 (GB)	裸芯容量 (Gb)	堆栈高度	HBM堆栈数量	发布日期	封装技术
	Inferentia 2	AI ASIC	HBM3	32	16	16	8	2	2022-Q4	CoWoS-R
	Trainium 2	AI ASIC	HBM3E	96	24	24	8	4	2022-Q4	CoWoS-R
	Inferentia 2.5	AI ASIC	HBM3	32	16	16	8	2	2022-Q2	CoWoS-?
	Trainium 2.5	AI ASIC	HBM3E	144	36	24	12	4	2022-Q2	CoWoS-?
	Trainium 3	AI ASIC	HBM3E	144	36	24	12	4	2022-Q4	CoWoS-S
	MI200	GPU	HBM4E	128	16	16	8	8	2021-Q4	FO-EB
	MI300	GPU	HBM3	192	24	16	12	8	2023-Q4	CoWoS-S
	MI325	GPU	HBM3E	288	36	24	12	8	2024-Q4	CoWoS-S
	MI355	GPU	HBM3E	288	36	24	12	8	2025-Q4	CoWoS-S
	MI400	GPU	HBM4	384	48	24	16	8	2026-Q1	CoWoS-?
	TPU v5e	AI ASIC	HBM4	16	16	16	8	1	2023-Q2	CoWoS-S
	TPU v5p	AI ASIC	HBM4E	96	16	16	8	6	2023-Q4	CoWoS-S
	TPU v6e	AI ASIC	HBM3E	32	16	16	8	2	2024-Q3	CoWoS-S
	TPU v6p	AI ASIC	HBM3E	192	24	24	8	8	2025-Q2	CoWoS-S
	TPU V7p	AI ASIC	HBM4	288	36	24	12	8	2025-Q4	CoWoS-?
	TPU v7e	AI ASIC	HBM4	216	36	24	12	6	2026-Q3	CoWoS-?
	Gaudi2	GPU	HBM4E	96	16	16	8	6	2022-Q2	CoWoS-S
	Gaudi3	GPU	HBM4E	128	16	16	8	8	2024-Q4	CoWoS-S
	Maia	AI ASIC	HBM3	64	16	16	8	4	2025-Q3	CoWoS-S
	A100	GPU	HBM4E	80	16	16	8	5	2020-Q3	CoWos-S
	H100	GPU	HBM3	80	16	16	8	5	2022-Q4	CoWos-S
	H200	GPU	HBM3	144	24	16	12	6	2024-Q2	CoWos-S
	B200/GB200	GPU	HBM3E	192	24	24	8	8	2025-Q1	CoWos-L
	B300/GB300	GPU	HBM3E	288	36	24	12	8	2025-Q3	CoWos-L
	GB300A	GPU	HBM3	144	36	24	12	4	2025-Q3	CoWos-L
	R100	GPU	HBM4	288	36	24	12	8	2026-Q1	CoWos-L
	Dojo	AI ASIC	HBM4E, HBM3	160	32	16	16	5	2023-Q3	InFO-SoW

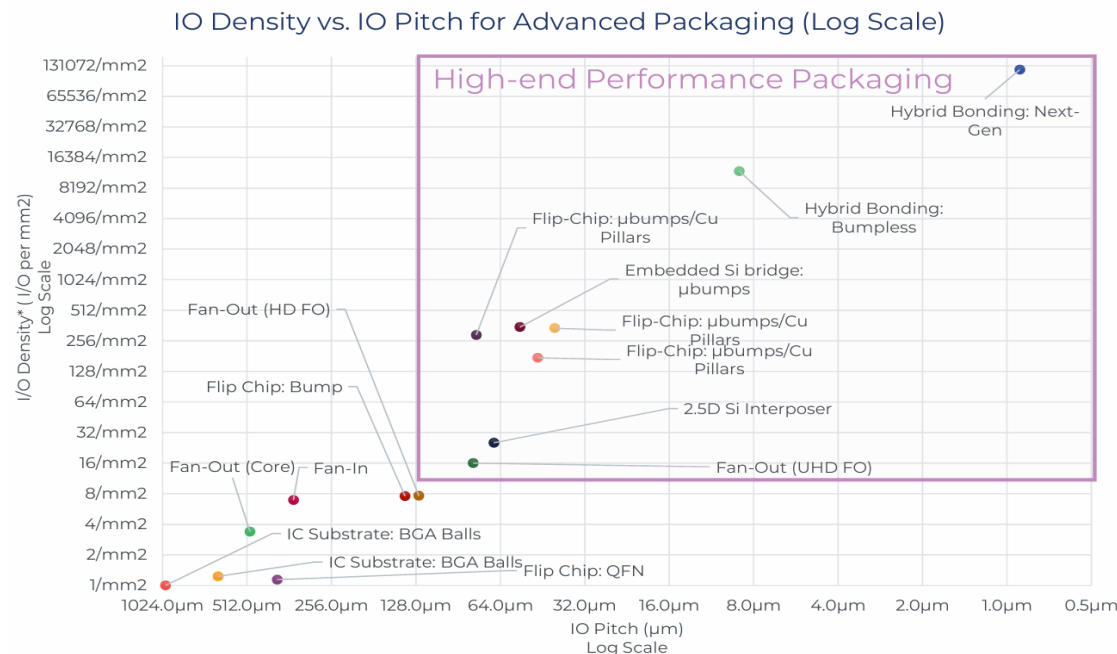
先进封装的摩尔定律：解互连

- **芯片间的互连与制程的差距。** 互连的作用是在封装内将一个成品半导体芯片与另一个连接起来，其核心目的是在芯片与印刷电路板之间快速、准确地传输电信号。更先进的封装技术通常伴随着封装尺寸和功耗的降低，以及互连密度（通常称为I/O数量）的提高。
- 当前封装中最常见的互连类型是引线键合，即通过极细的金属线将芯片连接到电路板上以传输电信号。然而，引线键合面临的挑战在于，其尺寸未能与晶体管密度的微缩保持同步，这意味着晶体管处理能力的增长已超出了引线所能传输的极限。
- 先进封装技术正试图通过采用新型互连方案来解决这一问题，例如使用“凸块”、“焊球”或“晶圆级封装”等方式替代引线来连接芯片，从而大幅提升其在芯片制造工艺的价值量，当前晶体管微缩的前端价值量远高于后端，而弥合晶体管处理能力增长与传输极限的“鸿沟”的关键在于先进封装，与此对应的也是先进封装

▶ 图表：前端晶体管微缩的价值量与后端的巨大差异，提升晶体管处理能力和传输差异的关键在于先进封装



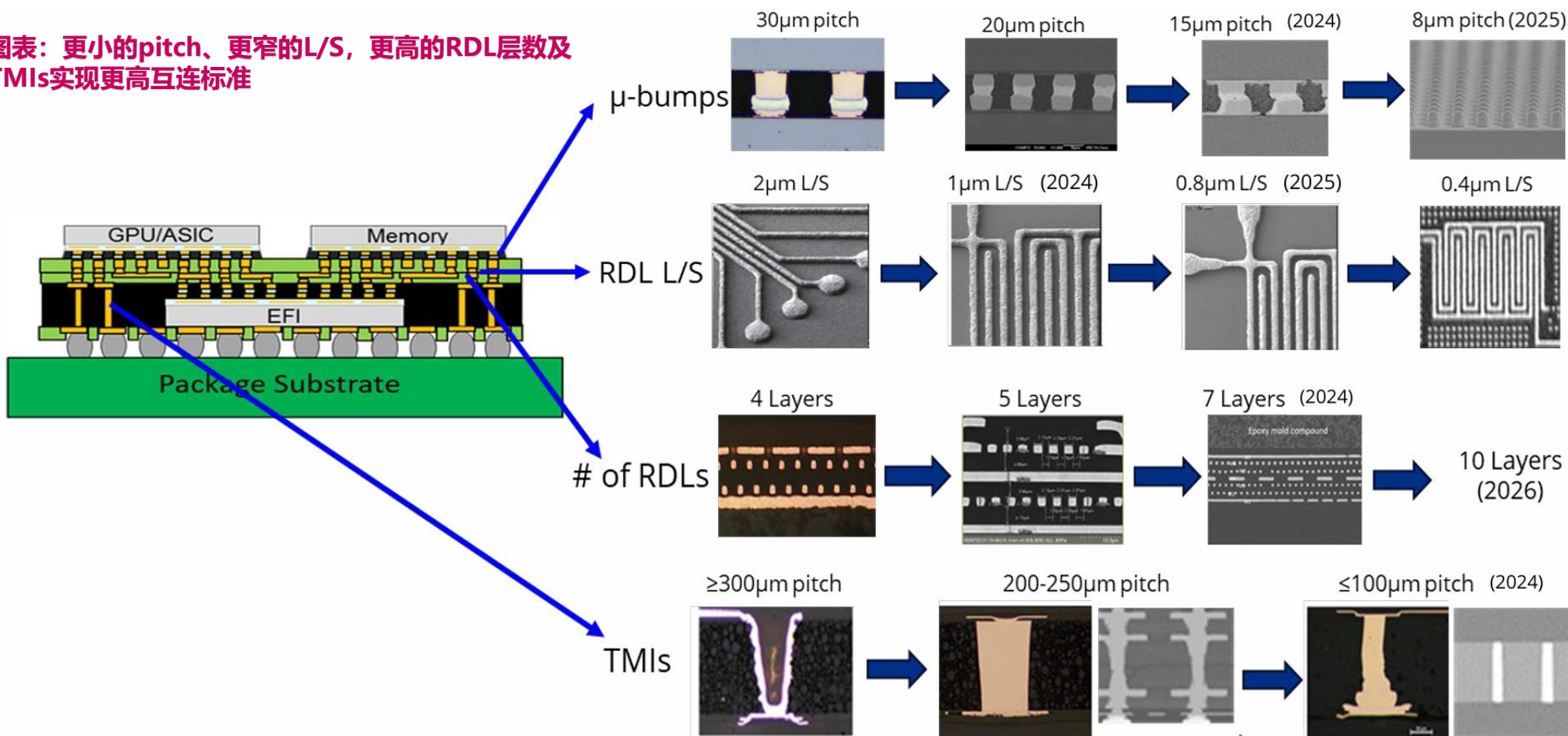
▶ 图表：通过封装工艺提升互连密度以实现晶体管处理能力与传输间的差异



先进封装的摩尔定律：解互连

通过缩小微凸块 (μ -bumps) 间距、提升RDL的布线密度 (L/S)、RDL层数及穿模通孔实现更高密度的电气连接

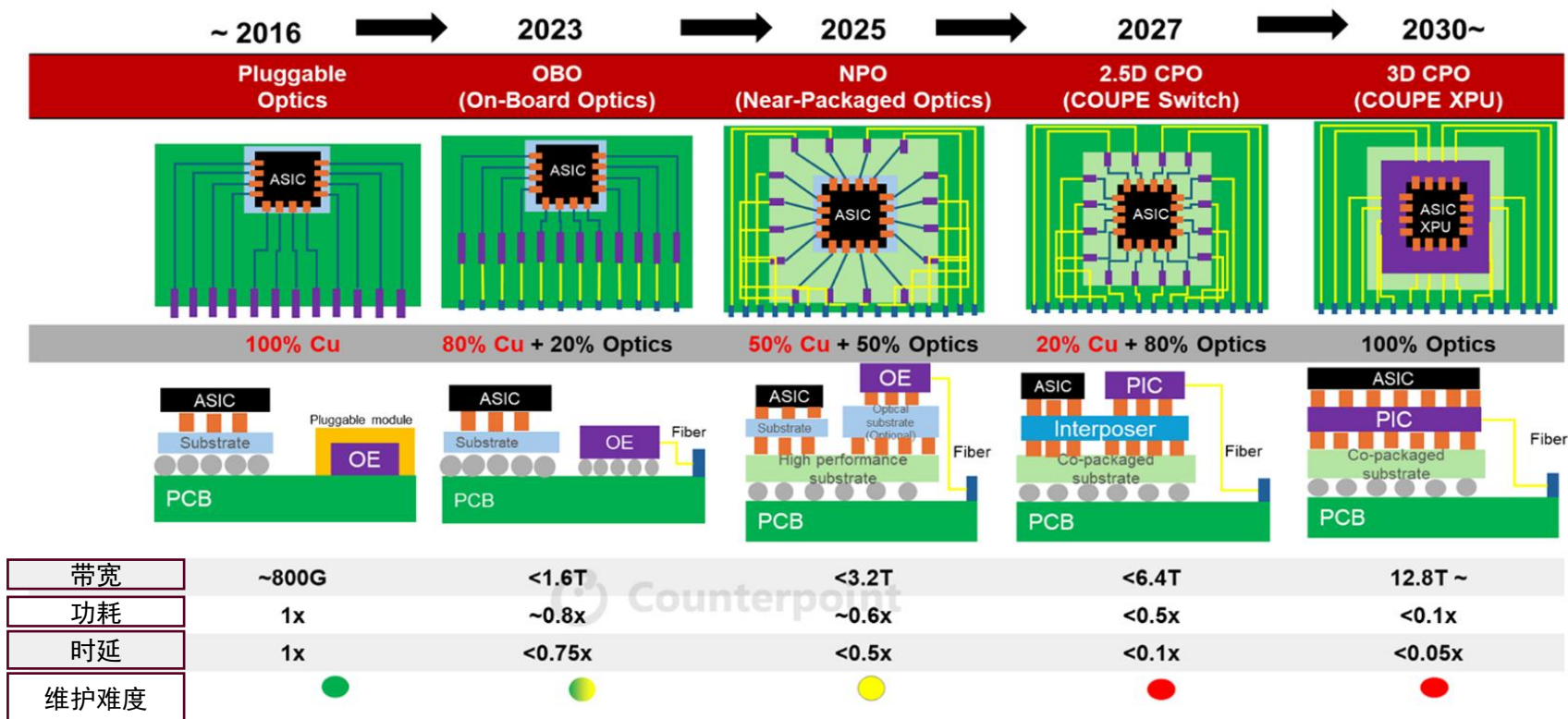
图表：更小的pitch、更窄的L/S，更高的RDL层数及TMI实现更高互连标准



先进封装的摩尔定律：解互连

- 此外，封装不仅局限于AI ASIC/GPU 封装体内部，随着机柜内互连及机柜间互连的要求提高，带宽及功耗问题也需要通过先进封装破局。共封装光学（Co-Packaged Optics, CPO）是将光学器件（如光引擎、收发模块）与大规模集成电路芯片（如交换ASIC或加速器芯片）封装在同一基板上的技术。
- CPO 的核心理念是利用异构集成技术，将处理光信号的光学引擎（OE）与处理电信号的逻辑芯片（如交换机 ASIC 或 GPU）封装在同一个基板上，以缩短电信号的传输距离，从而提高带宽并降低功耗。但是，2.5D及3D CPO面临着维护成本较高的问题。

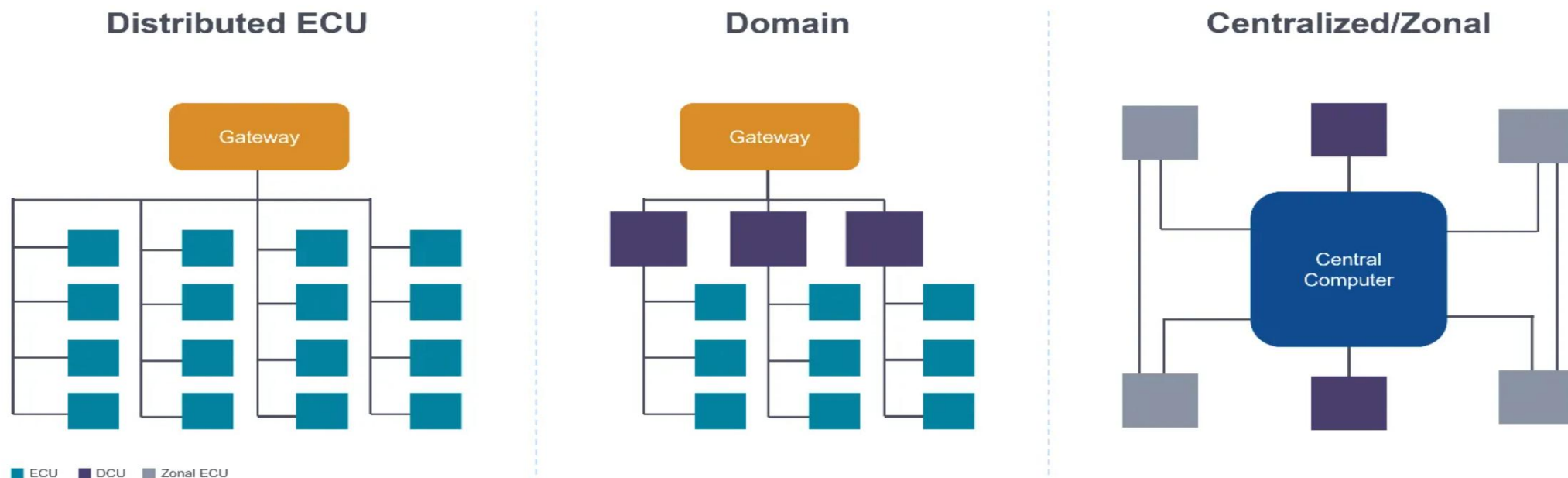
图表：通过光电共封装提升光电能效及降低时延



先进封装的摩尔定律：解边缘AI

- 在云端推理和边缘AI方面，先进封装带来的高带宽和小型化同样重要。许多推理ASIC需要在较小功耗下实现高速内存访问和灵活I/O，例如Habana、Groq等公司的推理芯片就通过MCM封装片上SRAM及I/O芯片，实现计算与存储分离但又紧耦合的结构。
- 边缘设备由于空间受限，更依赖系统级封装(SiP)整合多芯片：典型案例是Tesla FSD自动驾驶芯片，将NPU、GPU、CPU和高带宽DDR等通过FCBGA和倒装PoP封装在一起，既满足算力又控制体积。在自动驾驶计算方面，L4/L5级别自动驾驶需要的算力和带宽逼近数据中心水平。例如高级ADAS SoC将集成CPU、GPU/NPU以及多个高速图像处理单元，还需与外部DRAM高速通信。传统单芯片设计会因面积过大、良率过低而受限，先进封装提供了折衷方案，通过Chiplet拆分SoC并封装集成。

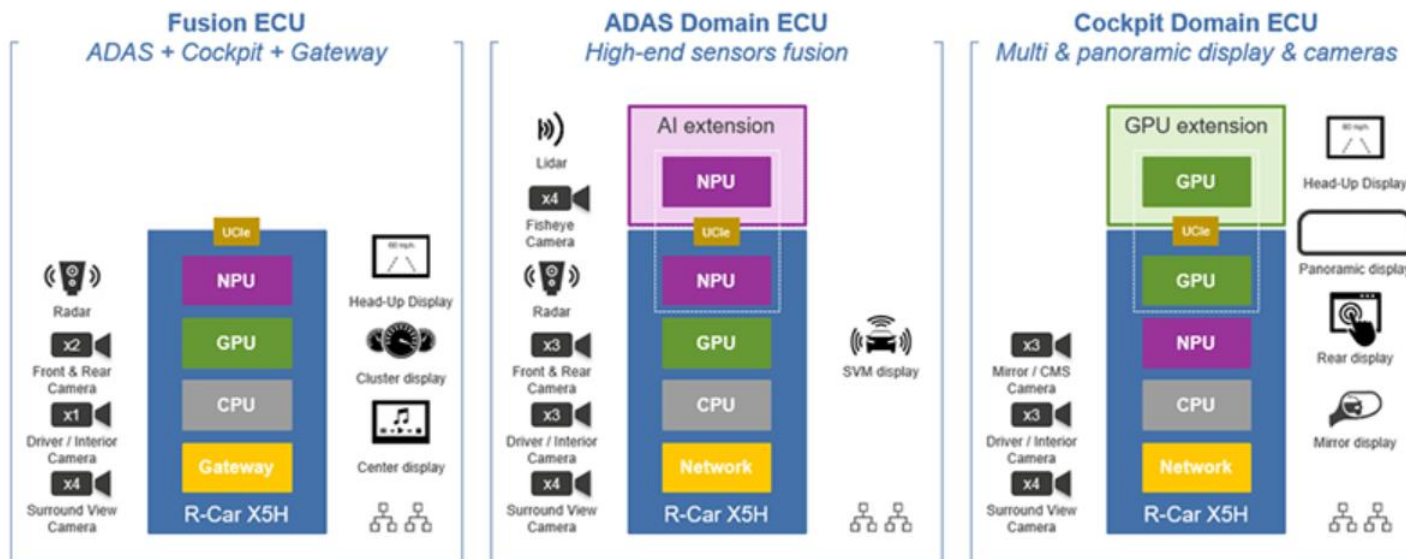
➤ 图表：先进封装通过 Chiplets 方案来实现空间受限下的高带宽、算力需求



先进封装的摩尔定律：解边缘AI

- **ADAS计算需求增长的两个关键因素：更多高分辨率传感及“集中化+AI”**：“更多传感”意味着电动车配备了更多及更高分辨率的传感器，导致大幅增加了需要处理的数据量；第二是“集中化+AI”，将车辆数据处理集中在少数运行AI算法的强大电子控制单元中。传感器数据的爆炸性增长与向集中式AI驱动处理的转变，共同推动了对ADAS系统更强计算能力的需求。
- ADAS工作负载所需的各种处理器类型包括用于常规控制任务和传感器管理的MCU，以及用于处理复杂算法的更强大的APU等应用/加速器芯片，专用的VPU用于摄像头和图像处理，其中多传感器VPU能够同时处理并融合来自多种传感器模式的数据。此外还有SoC FPGA，它结合了传统处理器和可重新配置的逻辑单元，以加速特定功能。最后，一个中央处理器将一切整合在一起——这是一个高性能处理器，负责集成和处理所有输入的传感器数据，并运行核心的驾驶辅助或自动驾驶软件。
- 由于复杂的处理任务，ADAS及车载芯片转向Chiplet及先进封装，瑞萨推出了R-Car X5H，其核心AI能力可通过多芯片封装的AI芯片组扩展，通过异构集成的方式满足不同场景需求。

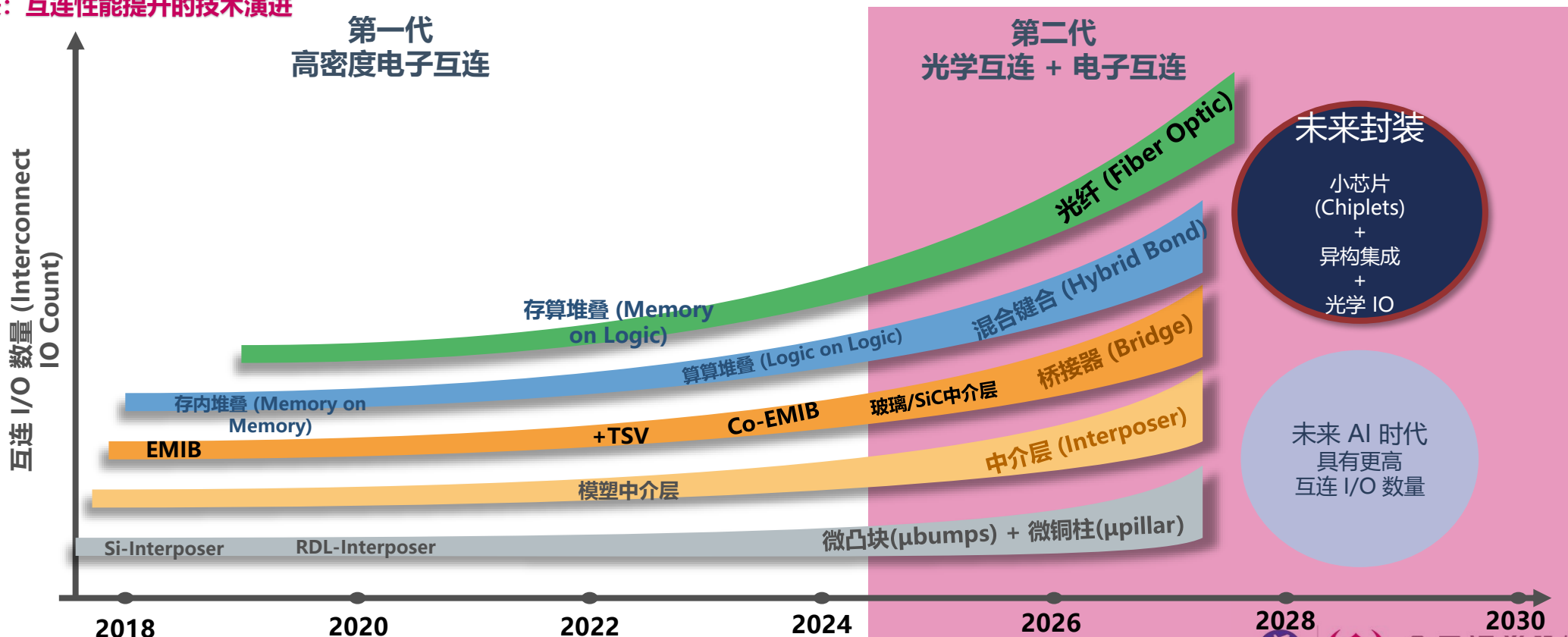
➤ 图表：通过Chiplet从而实现汽车电子不同场景需求



先进封装技术演进

- 先进封装的技术演进核心是互连 I/O 数量与带宽密度持续上升：第一代以高密度电子互连为主，从 Si-Interposer、RDL-Interposer 到 EMIB/Co-EMIB（叠加 TSV），再到中介层与桥接器等形态，在微凸块/微铜柱基础上实现从存内堆叠到存算堆叠与算算堆叠的更高集成；
- 随着间距继续缩小，混合键合（Hybrid Bond）成为提升互连密度与能效的关键。进入第二代，封装不再只通过电连接，而是把光互连引入封装体系（Fiber Optic + 电子互连），面向“未来封装 = 小芯片（Chiplets）+ 异构集成 + 光学 I/O”的方向演进，以支撑 AI 时代更高的互连 I/O 需求并缓解带宽与功耗瓶颈。

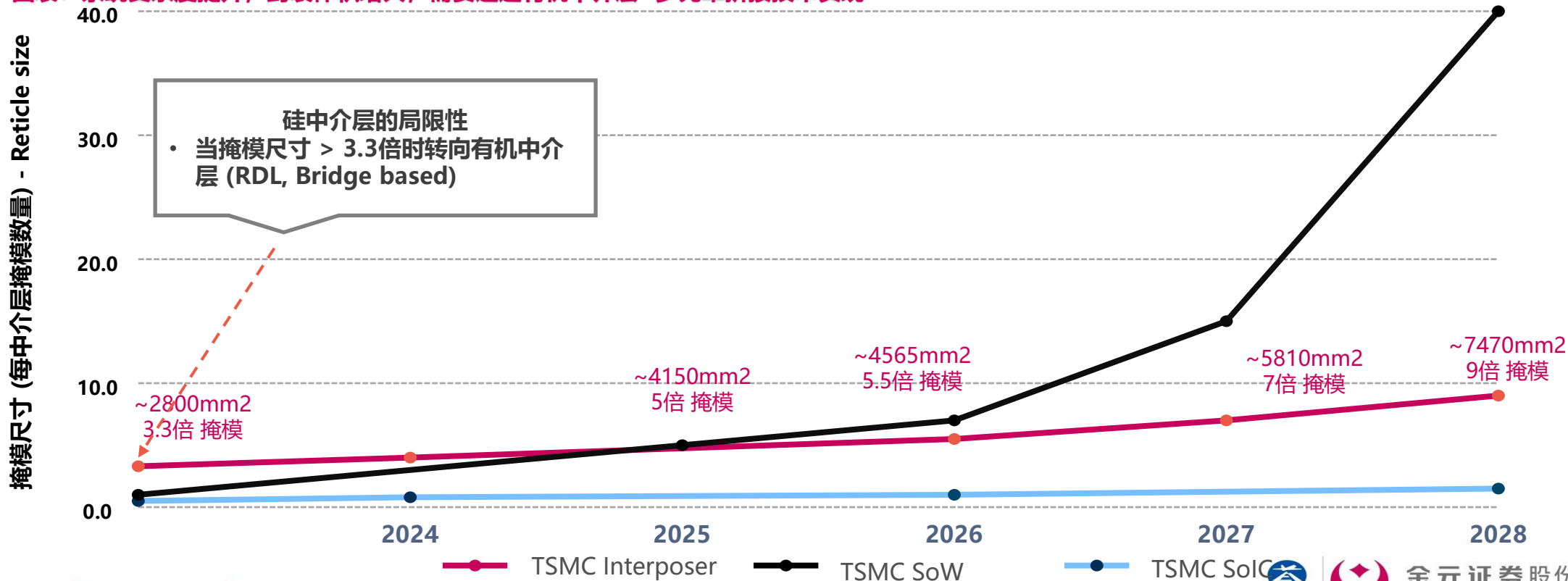
➢ 图表：互连性能提升的技术演进



先进封装技术演进

- 随着系统复杂度提升，封装体积增大，封装正在通过多光罩拼接 (multi-reticle stitching) 把可制造的最大芯片/中介层面积推到光刻光罩尺寸极限之外，从而获得更大的互连载体来把 GPU 与 HBM 等存储更紧密地集成在一起，满足生成式 AI 驱动的 HPC 对带宽与互连 I/O 的爆发式需求。
- 此外，传统硅中介层在约 3.3× reticle 规模附近开始受到限制。若要继续放大到 5×、7×乃至 9× reticle (面积从约 2800mm² 提升到约 7470mm²)，需要更多转向以 RDL/桥接为基础的有机中介层方案。同时 SoW、SoIC 等系统级封装形态也将借助拼接把系统面积做大，支撑更高集成度与更高带宽。

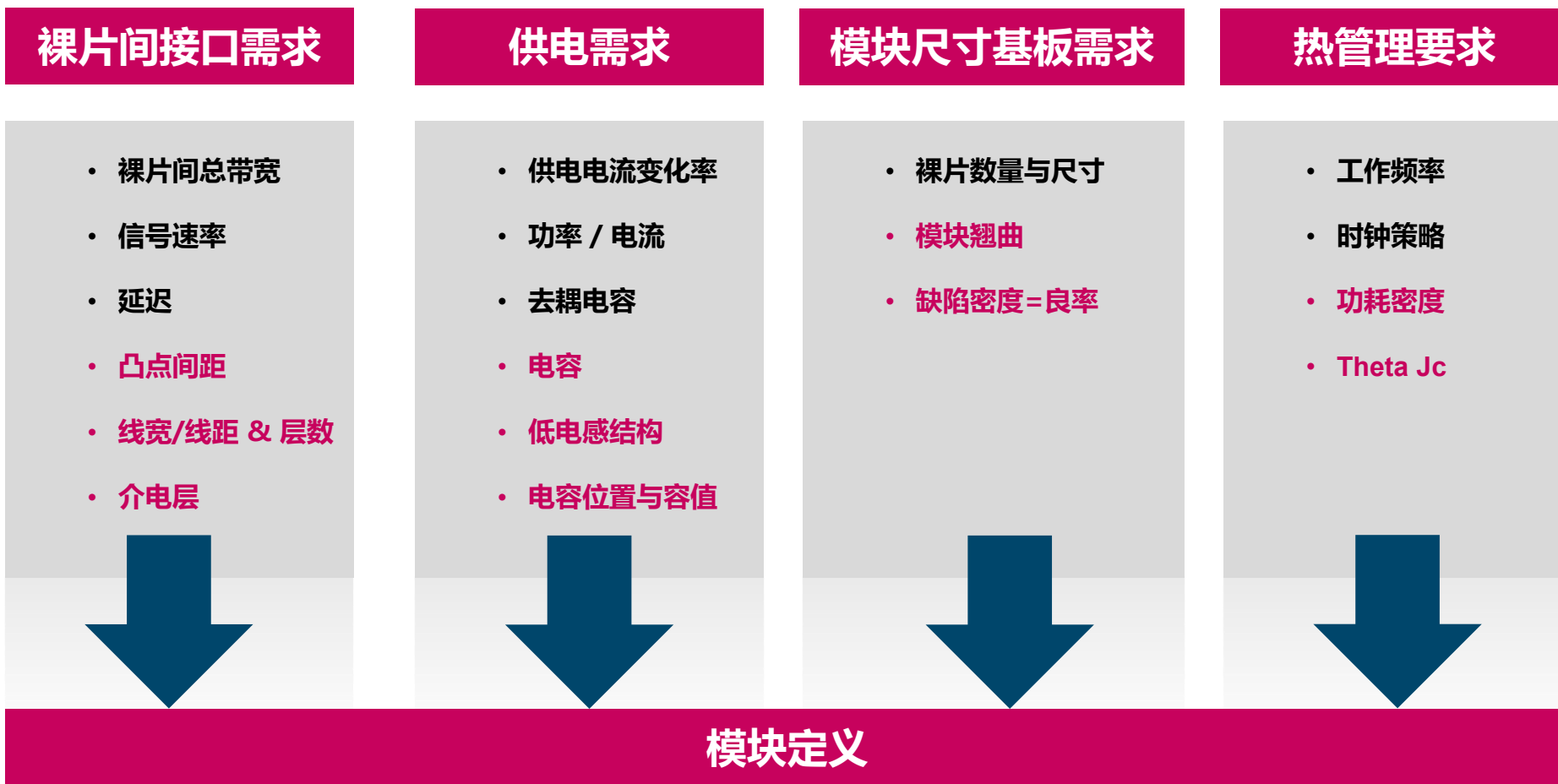
图表：系统复杂度提升，封装体积增大，需要通过有机中介层+多光罩拼接技术实现



先进封装技术演进

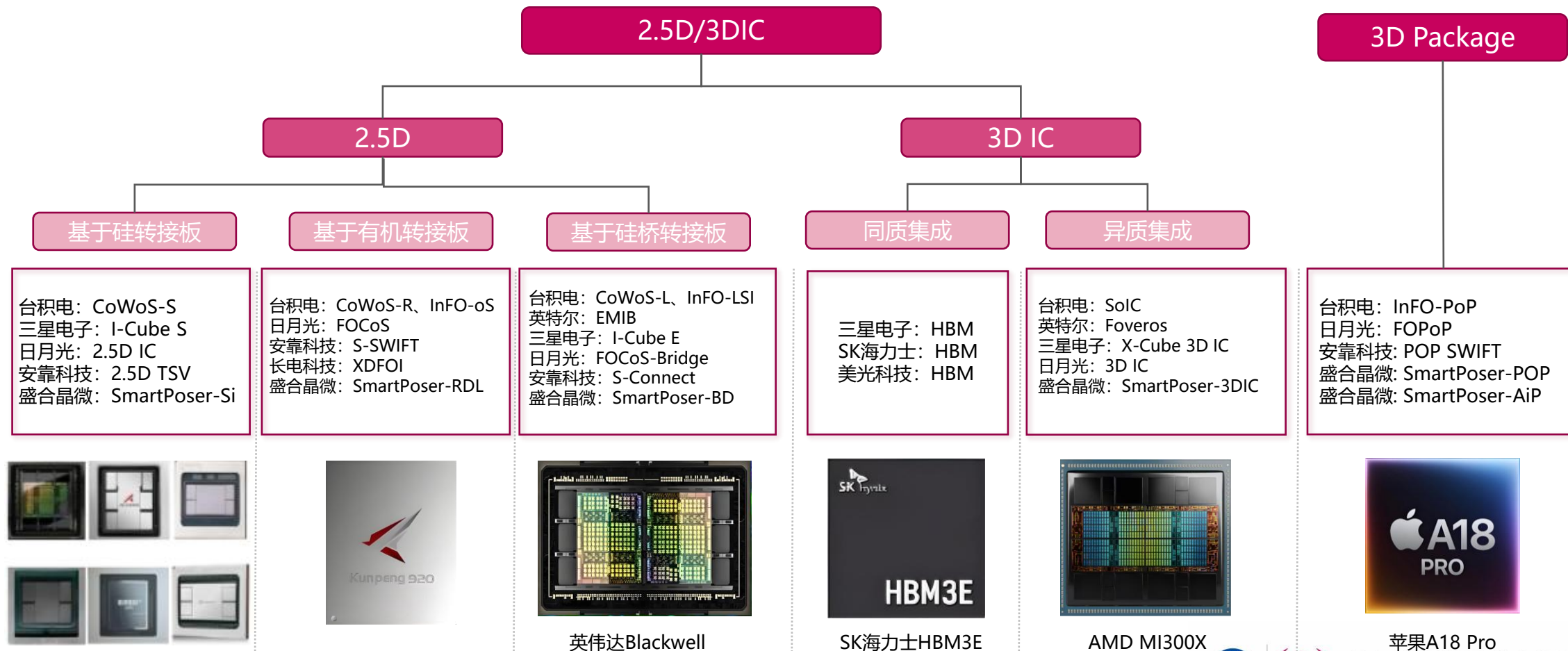
- 在设计和定义一个先进封装（如 Chiplets、异构集成模块）时，不存在一个单一的最优解，而是需要系统性地权衡多个相互关联、甚至相互制约的技术要求。
- 对于一个封装体（模块）的设计需要考虑四个维度：
 - **1、芯片间互连要求：**追求高性能。需要更高的总带宽、更快的线速和更低的延迟，但往往制造工艺需要更精细的凸点间距、更多的布线层数和更复杂的中介层；
 - **2、电源输送要求：**追求高稳定、高效率。需要应对更高的功率/电流，使用更多的去耦电容，并解决由高电流变化率 ($V=L di/dt$) 带来的电压噪声问题，可能会占用布局空间并增加设计复杂度；
 - **3、模块尺寸与基板要求：**追求高集成度与高良率。集成的芯片数量多、尺寸大，但会加剧模块翘曲，并因基板缺陷密度而影响最终良率；
 - **4、热管理要求：**追求高效能与可靠性。更高的工作频率和更激进的时钟策略会带来更高的功耗和功率密度。如果散热设计无法跟上，将导致芯片过热、性能下降甚至失效。

➤ 图表：封装设计需要考虑的因素



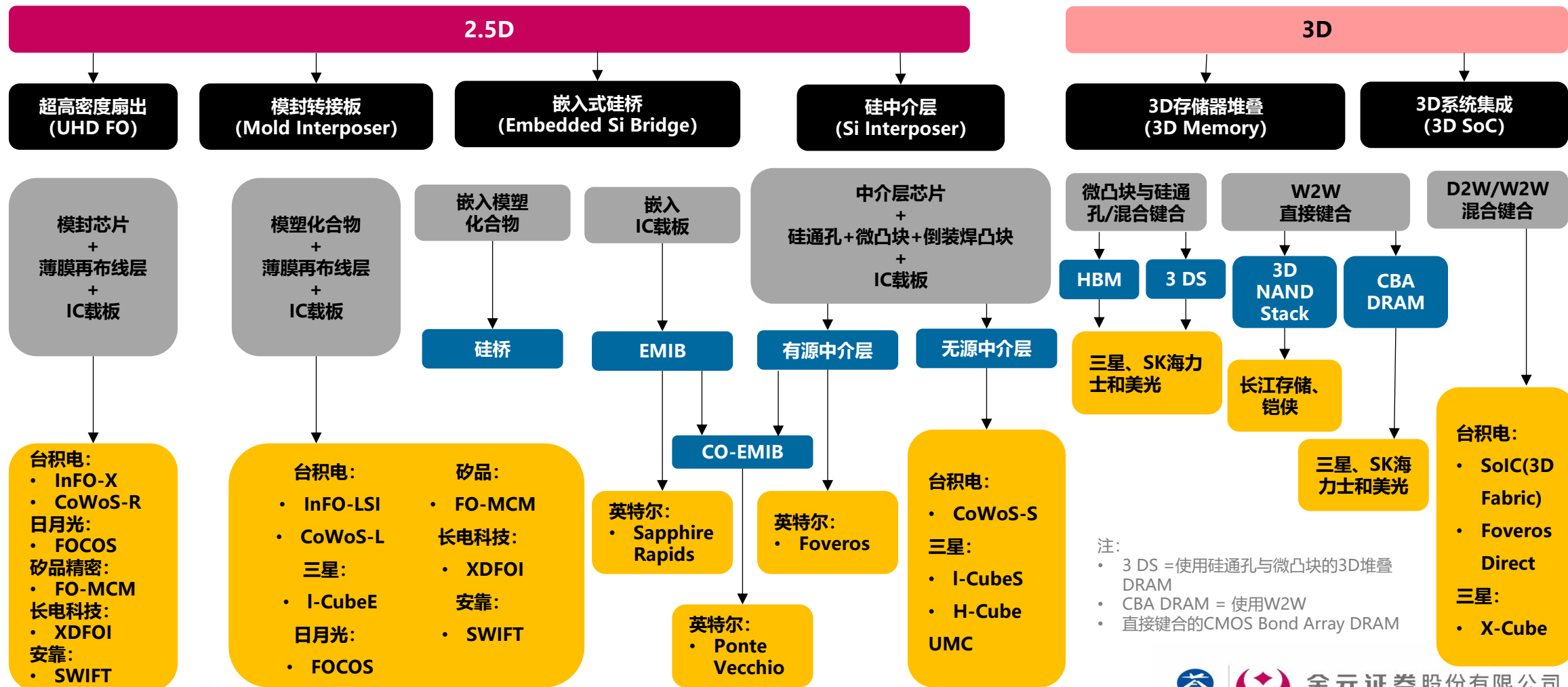
先进封装技术演进

- 先进封装可大致分为：2.5D、3D IC及 3D Package，3D IC与3D Package的区别在于前者强调裸片的直接集成，而后者更多是封装体之间的集成。
- 2.5D封装可基于中介层（转接板）的不同分为Si、有机以及硅桥。硅桥更像是介于硅转接板与有机转接板的“中间体”，其主要采用有机转接板，只在GPU和HBM显存之间需要高密度互连的局部区域，嵌入小块的硅片（LSI）。



先进封装技术演进

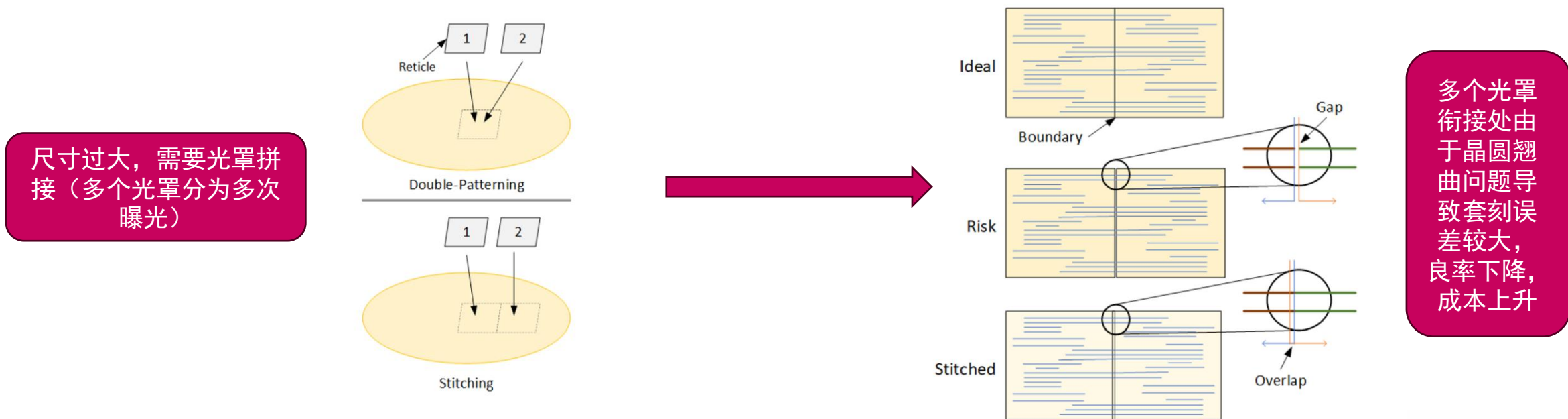
高性能封装技术



先进封装技术演进-2.5D

- 2.5D封装通过在中介层 (interposer) 上并置多个裸芯片，实现类似单芯片的互连密度，从而突破PCB/基板互连的线宽/线距限制。中介层最早以硅材料为载体 (硅中介层)。
- 然而硅中介层尺寸受限需要通过拼接实现光刻，但由于硅中介层面积大且薄，翘曲导致对准难度大，套刻误差增大限制良率从而导致成本较高。近年来业界积极探索多种替代方案，包括重布线中介层 (RDL Interposer)、有机模塑中介层 (Mold Interposer)、玻璃中介层 (Glass Interposer)、硅桥接 (Silicon Bridge) 以及面向光学I/O的光子中介层 (Photonic Interposer) 等。
- 不过，当前由于硅中介层的热膨胀系数与芯片相同，可减少热管理方面的挑战，因此仍继续用于高性能产品。此外，该技术已成熟，并获得了包括晶圆厂、IDM和OSAT在内的多家厂商的认证。

➤ 图表：当超过reticle尺寸时，需要通过光罩拼接 (stitching)，而在衔接处对准难度大，可能会导致良率下降



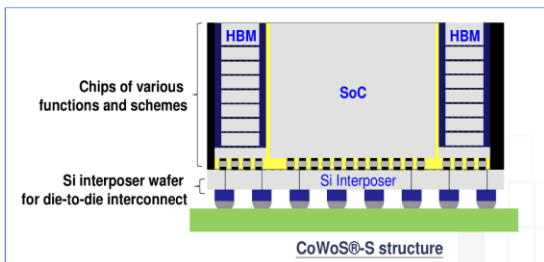
先进封装技术演进-2.5D

基于硅 (Si) 中介层

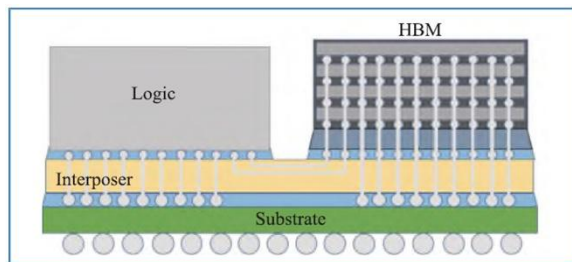
主要用于把 HBM 与逻辑处理器集成



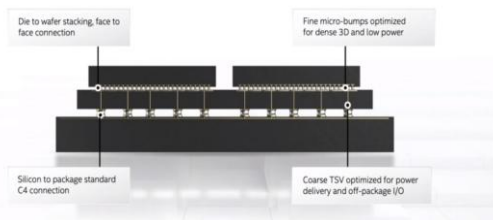
CoWoS-S



I-Cube



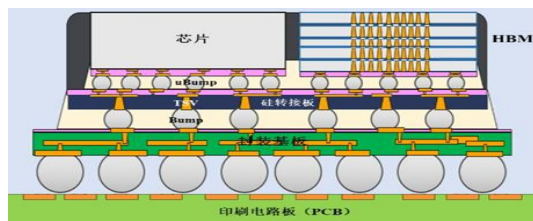
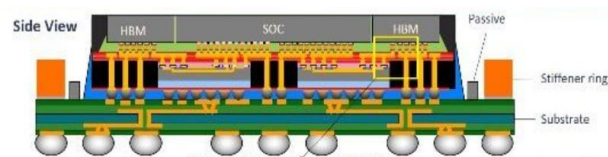
Foveros-S



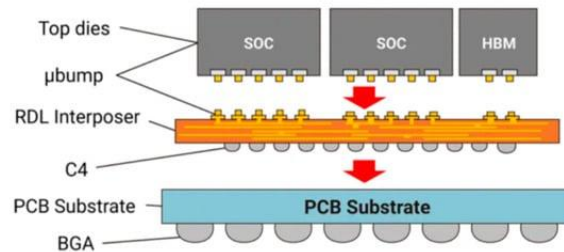
SmartPoser-Si



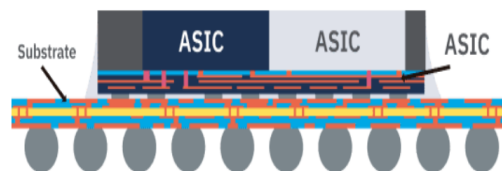
2.5D Package



CoWoS-R



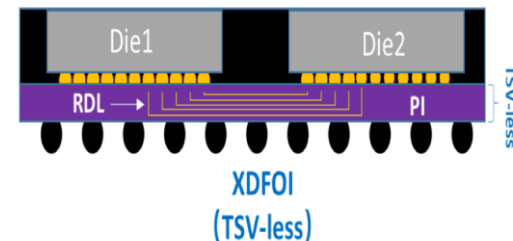
FoCoS-CF



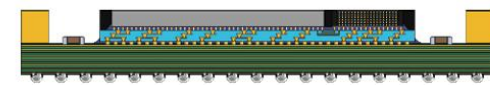
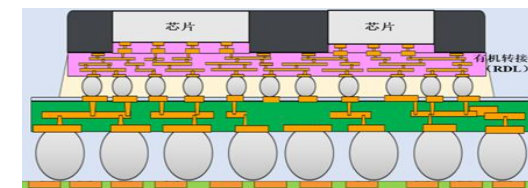
S-SWIFT



XDFOI-O



SmartPoser-RDL



Substrate SWIFT® Technology



www.jyzq.cn



全国统一客服电话: 95372

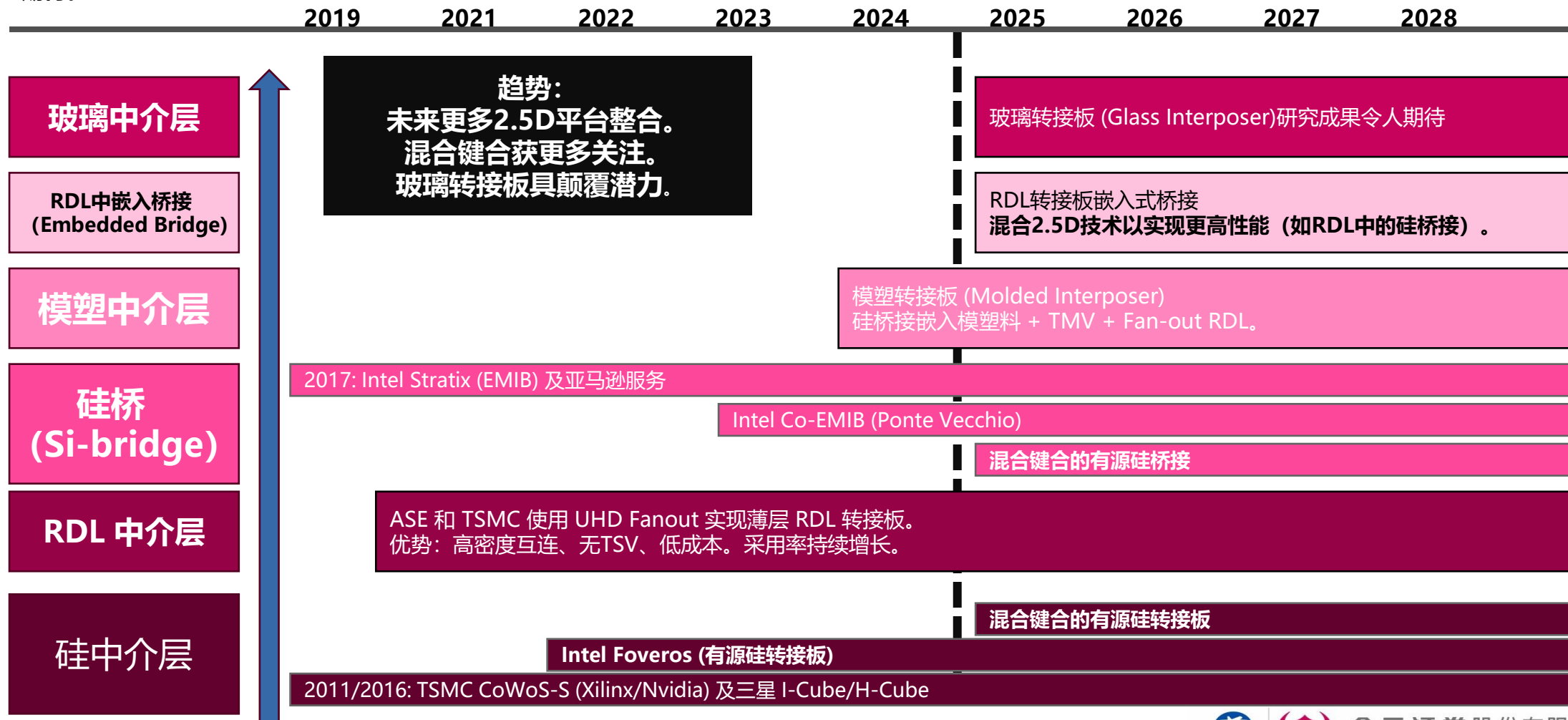
此文件版权归金元证券股份有限公司所有, 未经许可任何单位或个人不得复制、翻印。



金元证券股份有限公司
GOLDSTATE SECURITIES CO.,LTD.

先进封装技术演进-2.5D

- 2.5D封装趋势来看，由于硅中介层的尺寸限制，RDL及嵌入硅桥的有机中介层方案或成为主流。同时，2.5D+混合键合以及采用玻璃/SiC等新型材料中介层令人期待。



先进封装技术演进-2.5D

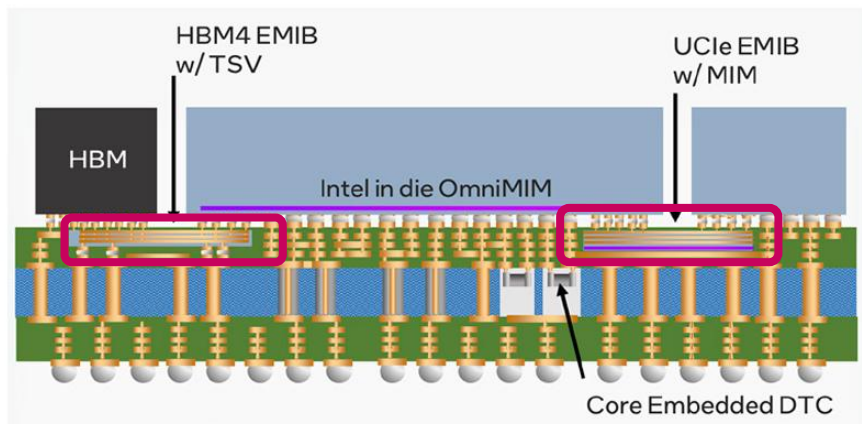
- 各类型中介层对比来看，硅中介层布线密度高+I/O Pitch小+采用TSV垂直互连且硅的介电常数更低，信号完整度更高，性能上更优，但是随着系统复杂度提升，硅中介层可扩展性及翘曲问题较为严重，成本极高。
- RDL中介层工艺相对简单且可扩展性较好，适用于成本敏感场景，但I/O密度及互连性能稍弱；模塑中介层（可在模塑材料嵌入硅桥及器件+TSV）可实现超过3.3×光罩尺寸的扩展，兼具性能与成本的平衡，或是未来 Chiplet 异质集成+高系统复杂度的首选。而IC载板内嵌入硅桥省略中介层，直接在基板上完成互连，由于IC载板布线密度与中介层差距较大，性能有下降。

参数	硅中介层 (Si Interposer)	RDL 中介层 (RDL Interposer)	模塑中介层 (Mold Interposer)	IC 载板内硅桥 (Si bridge in IC substrate)
成本	由于采用带 TSV 的大面积硅中介层，成本非常高	无 TSV (成本压力更小)	可在较小硅面积上引入 TSV	可在较小硅面积上引入 TSV
嵌入能力	可集成深沟槽电容 (DPC)	可在 RDL 中嵌入器件	可在模塑材料中嵌入器件 (复杂度更低)	可在载板中嵌入器件
集成能力	高	高	高	高
工艺复杂度	高, 需具备 FE (前端) 能力	中等	高, 工艺步骤多	中等
布线密度	精细金属布线 L/S	精细金属布线 L/S	精细金属布线 L/S	受 IC 载板限制, L/S 更大 (线宽/线距更粗)
I/O 间距 (Pitch)	在硅层可实现小间距, >30 μm 可行 (但昂贵)	小间距	在硅桥层可实现小间距, >30 μm 可行	小间距
信号完整性	良好	非常好	非常好	良好
是否需要 FE (前端) 能力	需要	不需要	需要 (用于硅桥), 但 TSV 可选	需要 (用于硅桥), 但 TSV 可选
OSAT 可实现性	仅适用于 CoW 和 oS 工艺	可以	可以 (硅桥晶圆除外)	可以 (硅桥晶圆除外)
翘曲问题	需要工艺控制	需要工艺控制	需要工艺控制	需要工艺控制
可扩展性	受光罩 (Reticle) 尺寸限制, 约 3.3×	可扩展至超过 3.3× 光罩尺寸	可扩展至超过 3.3× 光罩尺寸	可扩展至超过 3.3× 光罩尺寸

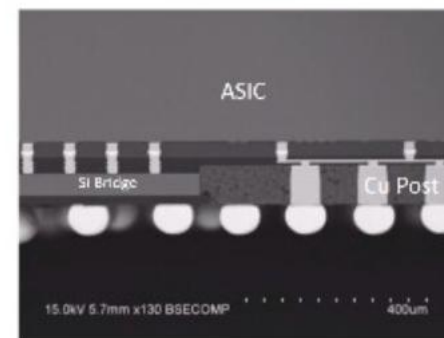
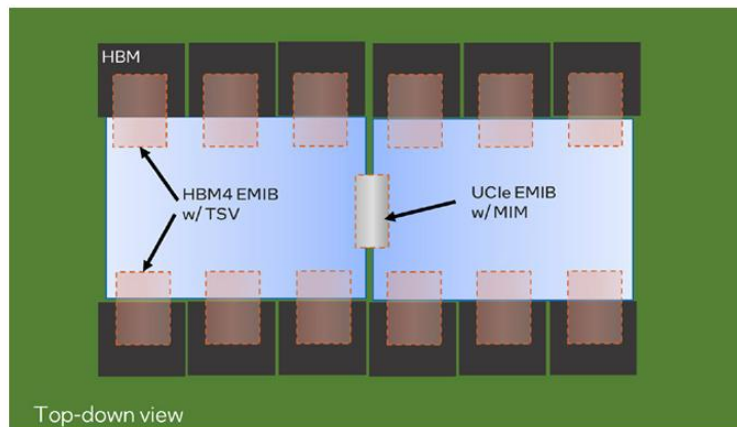
先进封装技术演进-2.5D

- 硅桥封装技术是一种2.5D解决方案，用于替代硅中介层技术。其核心是将一个或多个硅桥集成在特定的封装基板（可由不同材料制成）或模塑中介层中，以确保两个或多个芯片之间的互连。
- 英特尔率先开发了首项硅桥技术（即EMIB），将硅桥嵌入封装IC基板中。英特尔将该技术用于其Stratix 10 MX FPGA，随后又用于酷睿i7-8809G第八代处理器和Sapphire Rapids服务器处理器。此后，多家公司开始根据自身内部技术能力开发不同的解决方案。许多厂商开始研发将硅桥嵌入模塑料中并结合模塑料通孔和重布线层的方案。
- 桥接器数量取决于需要集成的芯粒总数，由于需异构集成更多芯粒和内存堆栈，封装中使用更多桥接器已成为趋势。例如英特尔 Sapphire Rapids 处理器存在多个硅桥接器。

➤ 图表：Intel EMIB，通过硅桥实现互连



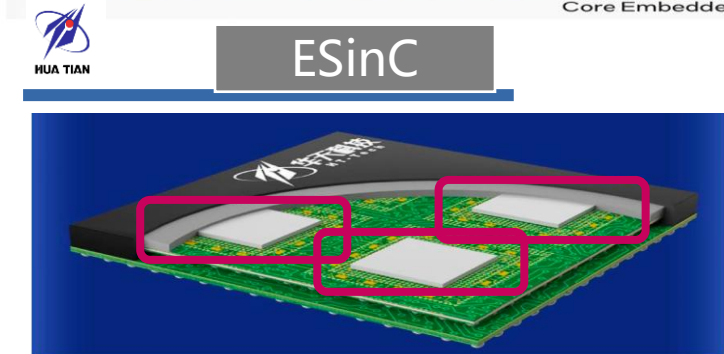
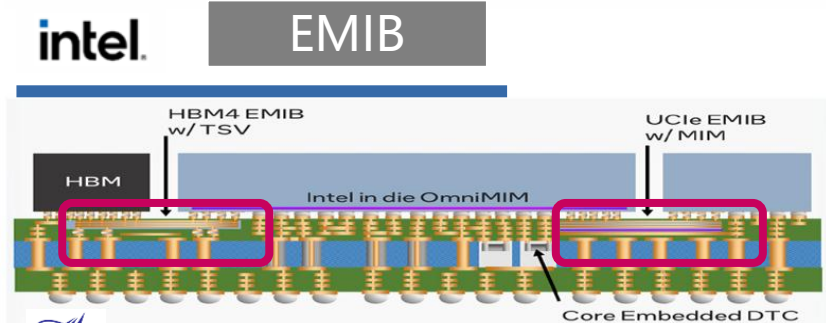
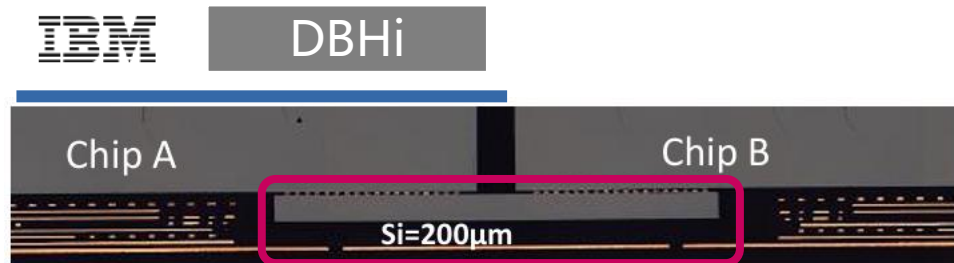
➤ 图表：SPIL 硅桥封装技术



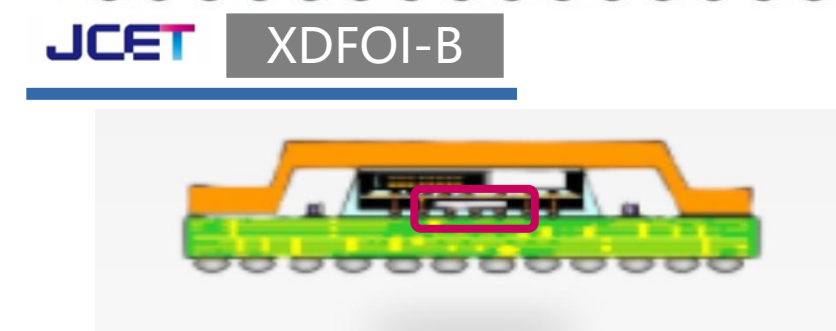
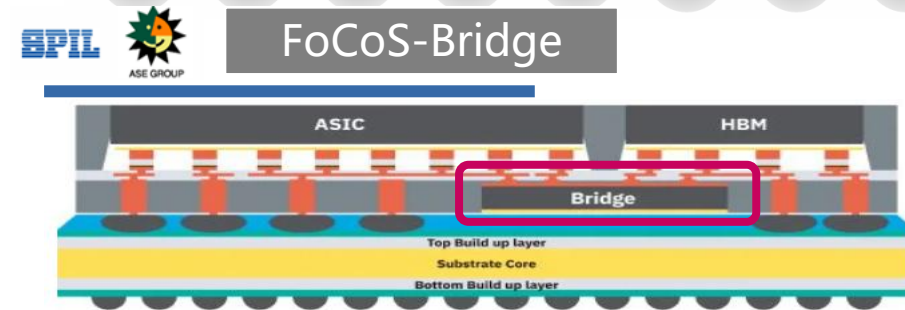
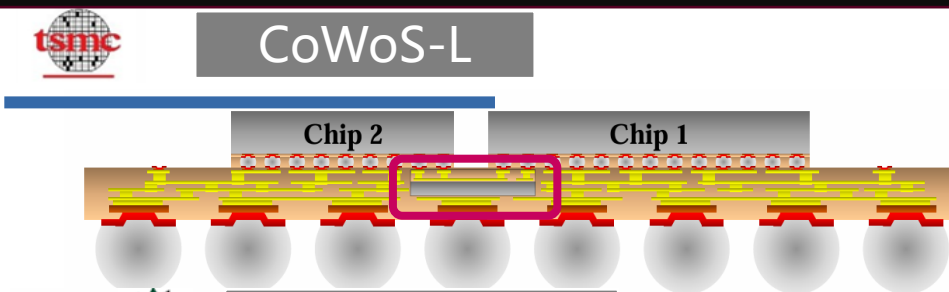
先进封装技术演进-2.5D

硅桥可以封装在基板上或内以及模塑体内，各公司工艺略有差异。嵌入式把硅桥放在基板腔体里再做布线，硅桥与基板过渡更像同一平面系统。而将硅桥封装在模塑体内，其布线密度要高于基板的方案。

硅桥封装在基板体内/上



硅桥封装在模塑体 (EMC) 内



















先进封装技术演进-2.5D

- 硅桥嵌入IC基板及硅桥嵌入模塑中介层对比来看，嵌入IC基板所需要硅桥面积较小，成本低于传统硅中介层，且工艺复杂度、成本略低于模塑中介层，但布线密度及I/O Pitch等互连性能不如模塑中介层方案。

方案	优势	局限
嵌入 IC 基板 (EMIB)	<ul style="list-style-type: none"> • 具备多芯片互连能力，可采用芯片后装 (chip-last) 方式实现逻辑芯片与 HBM 集成。 • 硅桥裸片可嵌入 IC 基板并进行测试；在后续装配前可进行 KGD (Known Good Die, 良品裸片) 筛选 • 与模塑中介层方案相比，工艺流程复杂度更低。 • 装配流程与 OSAT (外包半导体封装测试) 及 IC 基板供应商的能力体系兼容。 • 硅桥裸片尺寸在工程上几乎没有实际限制。 • 由于所需硅面积显著更小，硅桥的硅成本低于硅中介层 (Si interposer) 。 	<ul style="list-style-type: none"> • 虽然与硅中介层相比硅用量更小，但仍需在基板上开腔/挖槽 (cavity) 以放置硅桥，并通过粘接剂 (adhesive) 固定；相较传统倒装 (flip-chip) 会引入额外工序与成本。 • 与模塑体内嵌硅桥技术相比，EMIB 可实现的 I/O 间距 (pitch) 与走线尺度更大 (更粗)，互连密度相对受限。对于需要更小 I/O pitch 的应用，EMIB 无法完全替代硅中介层。 • 多层的大尺寸 IC 基板对热失配更敏感，易出现翘曲 (warpage)；需控制硅/IC 基板的比例以确保可靠性。
模塑中介层 (Mold Interposer)	<ul style="list-style-type: none"> • 具备多芯片互连能力，可采用芯片后装方式实现逻辑芯片与 HBM 集成。 • 在硅桥裸片层面可实现局部高密度互连，图形尺度可达亚微米级。 • 采用 RDL (Redistribution Layer, 重布线层) 互连，可获得比 IC 基板更精细的互连及更优的线宽/线距 (L/S) 缩放，从而改善信号传输。 • 装配流程与 OSAT 能力兼容。 • 硅桥裸片尺寸在工程上几乎没有实际限制。 • 硅桥集成方式灵活：可集成有源或无源硅桥，可带或不带 TSV (Through-Silicon Via, 硅通孔)。 • 可在模塑中介层内嵌入组件 (有源芯片、存储器、IPD (Integrated Passive Devices, 集成无源器件) 等)。 • 可对硅桥裸片与有源裸片进行 KGD 筛选。 • 由于所需硅面积显著更小，硅桥的硅成本低于硅中介层。 	<ul style="list-style-type: none"> • 中介层尺寸变大时，芯片—中介层以及中介层—基板之间的 CTE (热膨胀系数) 失配更显著，易引发翘曲 • 需进行材料 CTE 优化，以避免热失配导致的焊点疲劳、底填 (underfill) 脱层或开裂。 • RDL 可用面积受光罩版图视场 (reticle field) 限制，但业界正积极开发超视场 (larger-than-reticle) 中介层。 • 扇出重构晶圆 (fan-out reconstituted wafer) 可能出现芯片位移 (die shift) 问题。 • 装配流程复杂，包含多次芯片贴装步骤，成本高于传统倒装方案或 RDL 中介层。 • 硅桥晶圆仍需由晶圆代工厂 (foundry) 供货；若在硅桥上增加 TSV，将进一步提升工艺与供应链复杂度 (相关工艺需由代工厂管理)。

先进封装技术演进-2.5D

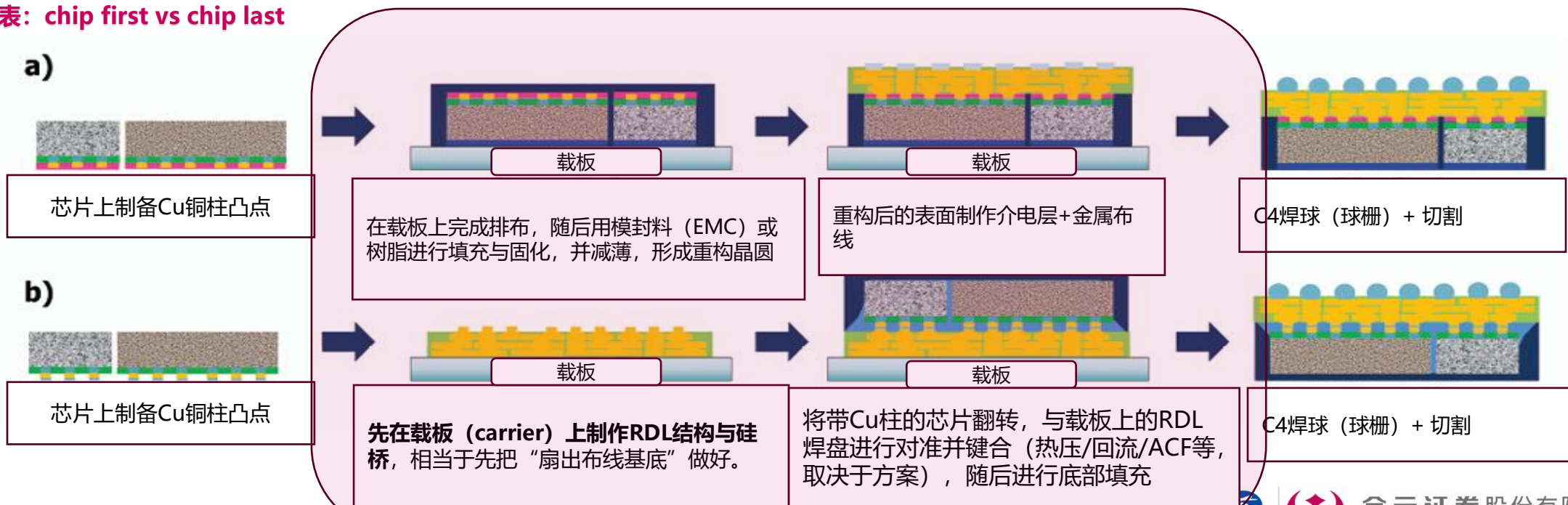
► 图表：主要公司硅桥封装的技术路线图

公司	技术	全称	Chip First	Chip Last or Middle	基板内/基板上	模塑封内	硅桥中进行TSV	状态	客户
 Intel	EMIB	Embedded Multi-die Interconnect Bridge			有机基板内		EMIB-T	量产	Intel, Amazon
 TSMC	Info-L	Integrated Fan-out with Local Silicon Interconnect						量产	Apple
	CoWoS-L	Chip-on-Wafer-on-Substrate - Local Silicon Interconnect						2025量产	Nvidia
 矽品精密	sFO-EB	Stack Fan-out embedded bridge					FO-EB-I	量产	AMD
 ASE	FOCoS-B	Fan-out chip on substrate-bridge						量产	AMD
 Amkor	S-Connect	Silicon Connect						小批量量产	
 Samsung	I-CubeE	Interposer Cube Embedded						未量产	
 长电科技	XDFOI-EB	X-Dimension Fan-Out Integration – Enhanced Bonding						小批量量产	中国客户
 盛合晶微	2.5D	2.5D					SmartPoser-BD	Si中介层量产/硅桥已验证	中国客户
 IBM	DBHi	Direct Bonded Heterogeneous Integration			有机基板上			未量产	
 PTI	PiFO	Pillar in Fan-out						研发	
 Nepes	Mold Int with Si Bridge	Mold Interposer with Si Bridge						研发	
 奕成科技	CM-MCM	Chip Middle Multi chip Module				?		2025	
 A-star IME	EFI	Embedded Fine Interconnect						研发	
 AOI	PSB	Pillar Suspended Bridge						研发	
 Mstech	FO bridge	Fan-out with bridge						未量产	
 华天科技	eSiNC	Embedded System in Chip			在硅基板内			研发	

先进封装技术演进-2.5D

- 在具体工艺上，2.5D封装可以分为chip first及chip middle/chip last。Chip-First 工艺将芯片先行放置在载板（胶带或玻璃载片）上，然后整体进行模压成型，形成重构晶圆，通过研磨暴露芯片表面并构建重分布层（RDL），完成铜凸块（C4）制备，然后去载板与基板互连。通常Chip First适用于少量芯片（约2-3片）、较小封装尺寸（ $<1\times$ 光罩）和低层数RDL（2~3层）
- Chip-Last 工艺先在载板上直接构建Fan-Out RDL基板（通常在玻璃或柔性材料上形成铜凸柱和RDL），然后将芯片倒装键合到RDL微凸点上并灌注环氧填充（Underfill）保护；之后对整个模块进行模压覆盖；最后去载板，在RDL背面制备C4凸块与PCB或基板连接。Chip Last 适合更大尺寸、多芯片的封装（如ASIC与多片HBM），可叠加多层RDL实现更高密度的互连性能。**
- Chip First的优势在于技术成熟，成本相对低，封装整体厚度可以很薄，但是模压封装后难以补救RDL缺陷，RDL良率问题可能会导致已封装芯片报废，并且芯片模压过程容易发生移位和翘曲问题，限制线宽和层数扩展。而Chip Last可以在RDL制程前先检查良率，不会损失KGD芯片，且更容易实现低L/S，键合后填充Underfill作为缓冲，但是对芯片与RDL对准精度要求高，成本较高

图表：chip first vs chip last



先进封装技术演进-2.5D

国内2.5D及嵌入封装技术也有一定突破，其中长电科技XDFOI解决方案覆盖有机转接板（RDL）及硅桥方案。盛合晶微在2.5D封装方案中技术较为领先，在硅中介层、RDL中介层以及硅桥中层（+TSV）均有覆盖，其中RDL、硅中介层已实现量产，并且具备Bumping、RDL等中道晶圆加工工艺能力。其他厂商包括通富微电、长电科技、甬矽电子均在开发2.5D封装方案。

图表：中国企业2.5D封装技术



长电科技提供XDFOI解决方案以集成芯粒，可选方案包括：

- XDFOI-O（有机转接板）
- XDFOI-B（硅桥）

其他解决方案包括：

- 硅中介层中的深沟槽电容
- 基板中的嵌入式电容
- 封装内的集成电压调节器（IVR）

已实现GPO集成PIC、EIC和ASIC用于高性能计算（HPC）系统。

华天科技 2.5D 技术：

- SiCS — 硅中介层芯粒系统
- FoCS — 扇出芯粒系统
- BICS — 桥联芯粒系统

华天嵌入式技术：

- eSiFO — 嵌入式硅基扇出封装（2.5D 解决方案）
- eSINC — 嵌入式芯片内系统（3D 解决方案）

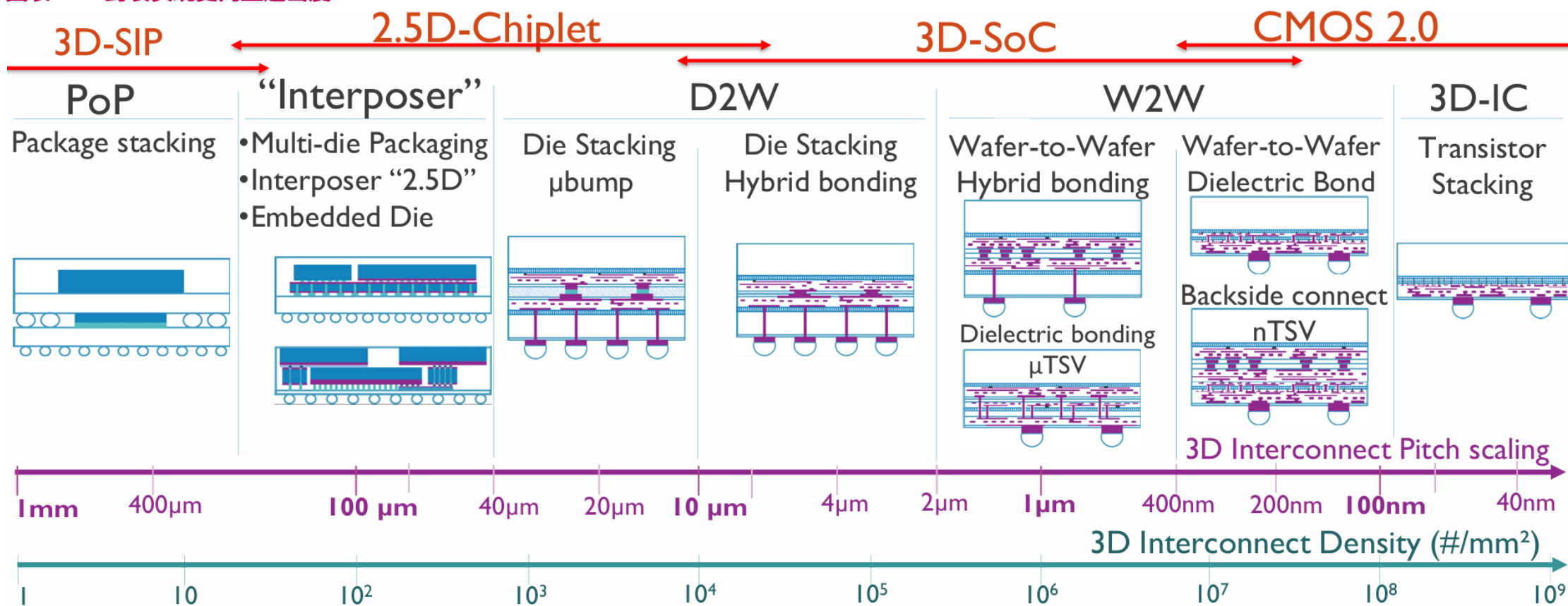
盛合晶微：提供与台积电 CoWoS-S 类似的 2.5D 技术，覆盖硅中介层、嵌入硅桥中介层、以及RDL中介层。主要客户为华为昇腾产品。

通富微电、甬矽电子与深科技已制定 2.5D 战略，正在开发对标台积电 CoWoS 的同类解决方案，从而向客户提供服务。

先进封装技术演进-3D

➤ 更高互连密度、更低功耗、更小封装尺寸的需求使得3D封装成为未来的大方向。即使使用2.5D封装，芯片间与芯片外部互连距离仍然受限。此外，2.5D封装扇出的互连密度受限于焊球，而3D封装通过铜柱可实现低于10微米的互连间距，且垂直堆叠带来了延迟和功耗优势。

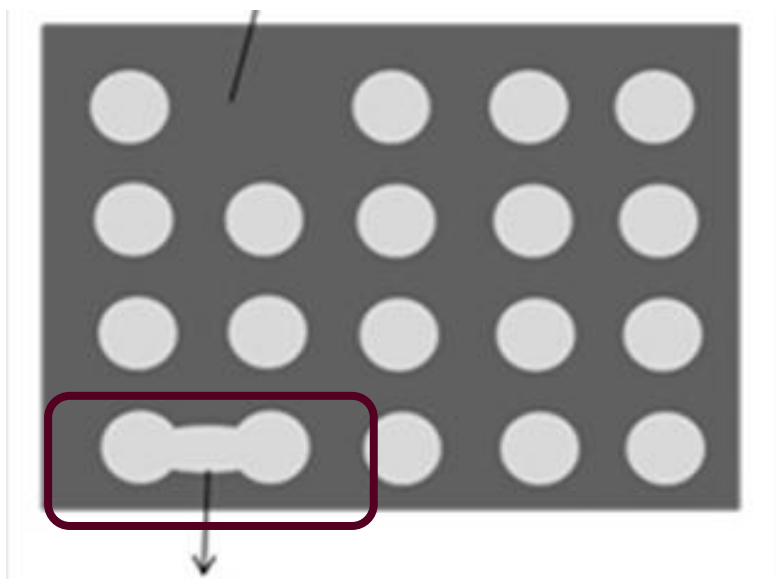
➤ 图表：3D封装实现更高互连密度



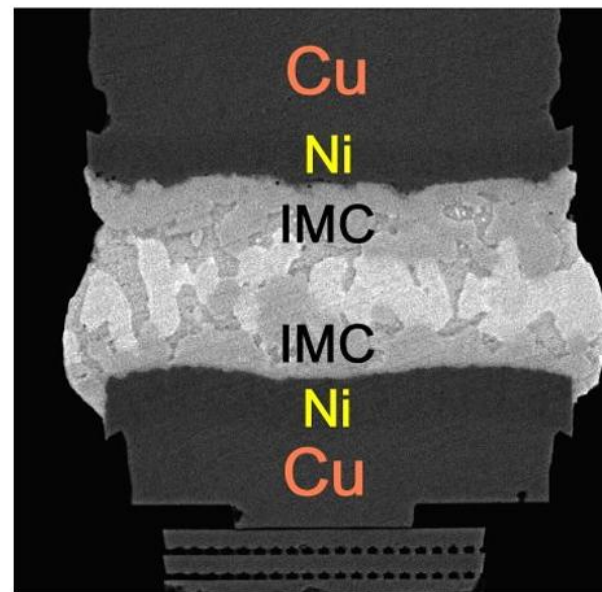
先进封装技术演进-3D

- **倒装芯片 (Flip Chip) 技术通过受控塌陷芯片连接 (C4) 凸块支撑了高性能计算的发展。**然而，当互连节距 (Pitch) 缩小至 $10\mu\text{m}$ - $20\mu\text{m}$ 区间时，传统的微凸块 (Microbump) 技术遭遇了难以逾越的“互连密度墙”。
- 制约2.5D互连密度的主要由三个因素构成：首先是**焊料桥接 (Solder Bridging) 风险**。在回流焊过程中，熔化的焊料在表面张力作用下极易在狭窄的间距内发生短路，限制了凸块间距的进一步缩小；其次是**金属间化合物 (Intermetallic Compounds, IMC) 的生成**，随着焊料球体积的缩小，高电阻率且脆性的IMC在互连结构中的占比显著增加，导致信号路径的电阻急剧上升，严重恶化了电源完整性 (Power Integrity) 和信号完整性 (Signal Integrity)；最后是**底部填充 (Underfill) 工艺的挑战**，在极窄的芯片间隙中，毛细管力难以驱动填充胶无空洞地流动，导致机械可靠性下降

➢ 图表：焊料桥接问题



➢ 图表：随着凸点变小，IMC占比增大，电阻率上行

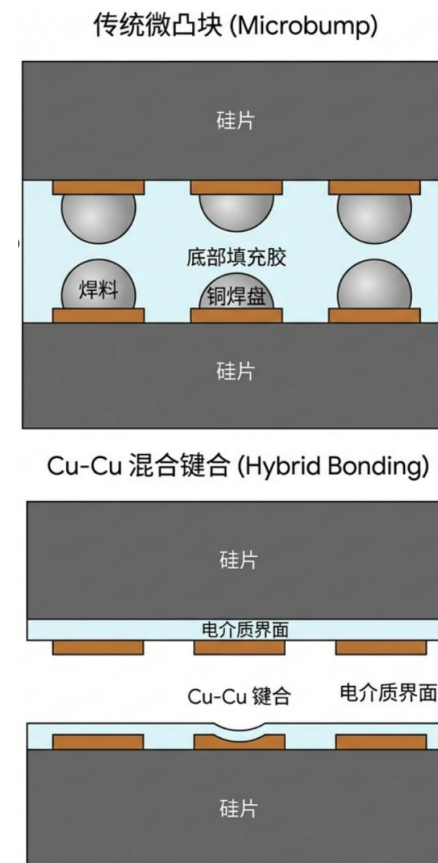


先进封装技术演进-3D

➢ 通过直接键合 (Direct Bonding) 和混合键合 (Hybrid Bonding) 实现“去焊料化”从而实现互连密度提升是3D封装的关键。混合键合技术通过在原子尺度上实现电介质与金属的直接连接, 消除了焊料层, 从而将互连间距从微米级 (20μm) 推升至<10μm。

➢ 图表: 混合键合实现高密度垂直互连

	焊料凸点 Solder Bump	铜柱 Cu Pillar	微凸点 μBump	混合键合 Hybrid Bond
凸点间距(μm)	>130	90 - 130	20 - 50	<10
凸点高度(μm)	>80	40 - 60	10 - 25	-0.015* - 0
直径(μm)	>100	45 - 65	10 - 25	<5
缺陷敏感度(μm)	>10	5 - 10	1 - 5	0.15 - 0.5
对准误差(μm)	>10	5 - 10	1 - 2	0.05 - 0.25
测量/计量精度(μm)	5	2	1	0.001



先进封装技术演进-3D

- 键合可以实现3D晶圆堆叠、亚10微米超细间距以及无凸块设计，实现了极高的I/O密度、带宽和存储密度，从而带来显著的系统性能提升与封装高度缩减；然而，其大规模量产仍受制于多重技术瓶颈，包括对表面平坦化与洁净度的苛刻要求、互连对准的高精度门槛、高温退火可能引发的器件损伤，以及目前在吞吐量、良率控制、电气测试标准和纳米级计量检测方面存在的局限性。

➤ 图表：键合工艺的优势及挑战

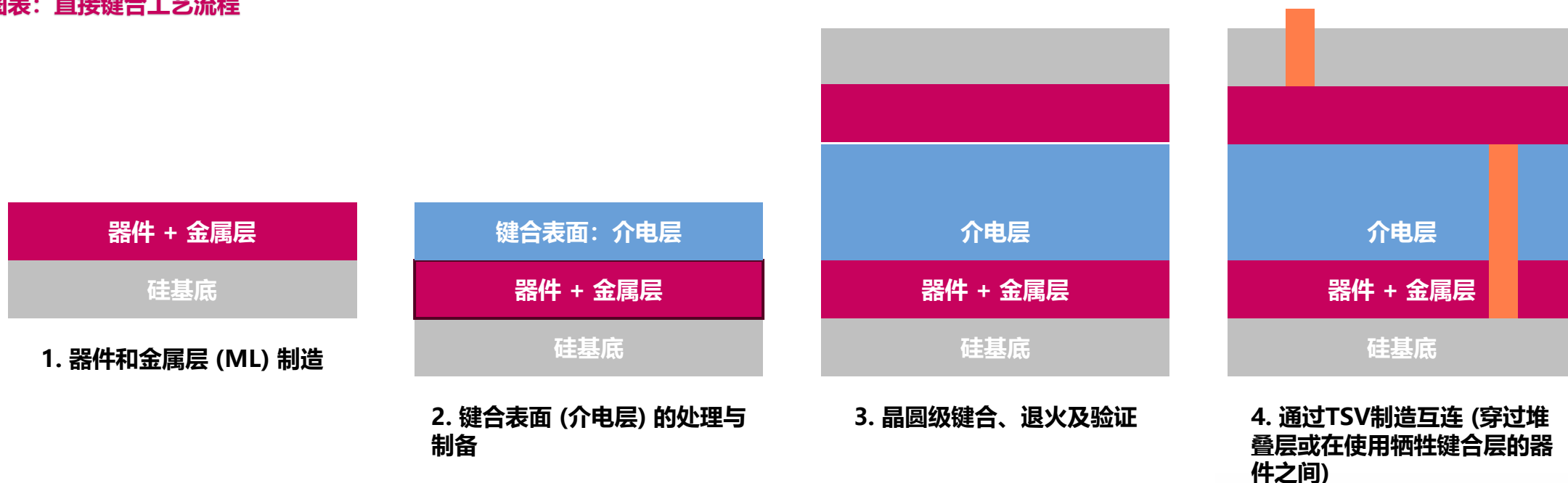
优势		挑战	
先进的3D晶圆堆叠	允许三维堆叠，实现更高性能和系统设计灵活性。	表面平坦化与洁净度	污染物会降低良率；需极高的表面洁净度以保证键合强度。
最大化 I/O 密度	实现更快的数据传输速率，提升系统整体性能。	互连重叠 (对准)	错位导致连接不良；对低接触电阻至关重要。
< 10μm 键合间距	实现亚10微米间距，允许集成更微小的组件。	高退火温度	200°C 对敏感的高端逻辑/存储器件来说仍然过高。
消除凸块需求	无信号损耗，提升性能，并显著降低封装高度。	吞吐量与良率受限	键合机能力限制了大规模生产的速度和良率。
更高的存储密度	集成高密度存储器，在更小空间内存储更多数据。	键合前后电气测试	测试困难，且目前尚无统一的测试标准。
高带宽，低延迟	改善信号传输质量，扩展带宽。	计量与检测	纳米级测量（间距、粗糙度、平整度）难度大，检测复杂。



先进封装技术演进-3D

- 直接键合，通常指不使用任何中间粘合剂，仅依靠分子间作用力将两个镜面抛光的表面结合在一起的技术。通常涉及电介质-电介质（Dielectric-to-Dielectric）的键合，最常见的材料是氧化硅（SiO₂）或碳氮化硅（SiCN）。
- 一般直接键合工艺包括表面活化与亲水性、退火、及TSV。晶圆表面必须经过极其严格的**化学机械抛光（CMP）**，使其粗糙度（Ra）降低到0.5纳米以下。随后，通过等离子体（Plasma）处理进行表面活化。等离子体轰击会打断表面的Si-O键，形成高活性的悬挂键，并使其极易吸附环境中的水分；为了获得高强度的永久键合，必须进行热退火处理。通常在200°C至400°C的温度下，界面处发生脱水缩合反应。

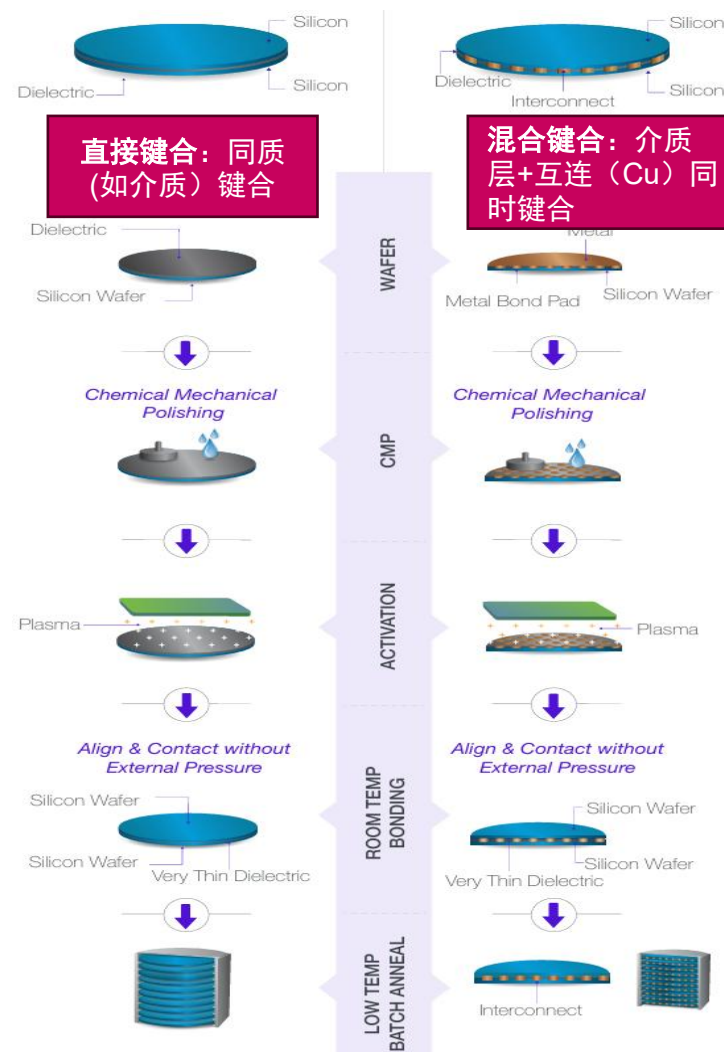
➤ 图表：直接键合工艺流程



先进封装技术演进-3D

- 混合键合 (Hybrid Bonding) 是直接键合技术的进阶形态。它在一个单一的界面上，同时实现了电介质的机械键合和嵌入式金属（通常是铜）的电气键合。
- 以adeia的“ZiBond Technology”和“DBI Process”对比来看，直接键合将同质，如介质层直接键合，形成介质-介质的键合界面，而混合键合是将介质层与互连层同时键合，形成垂直电通路。
- 对于直接键合，工序包括：**介质层制备** → **CMP** → **活化** → **室温键合** → **低温退火**，关键点在于洁净度、颗粒、表面粗糙度/平坦度、翘曲。
- 对于混合键合，工序额外需要实现**金属互连结构制作**。制作金属焊盘/互连（常见 Cu，也可 Ni），并与介质层共同构成键合面。（实际工艺里通常涉及沉积/电镀、刻蚀、阻挡/种子层等）
- 并且混合键合要求更高的平整度，金属焊盘与周围介质几乎同一平面（共面），否则会出现介质先顶住导致金属接触不上、或局部短路/空洞。
- 对于金属表面状态管理、对准、退火工艺，混合键合均有更高要求。因为要把两边的金属焊盘阵列一一对上（pitch 越小，对准越难），且退火工艺需要促进金属接触更充分（扩散/原子接触增强），最终得到稳定的互连。

➤ 图表：直接键合与混合键合



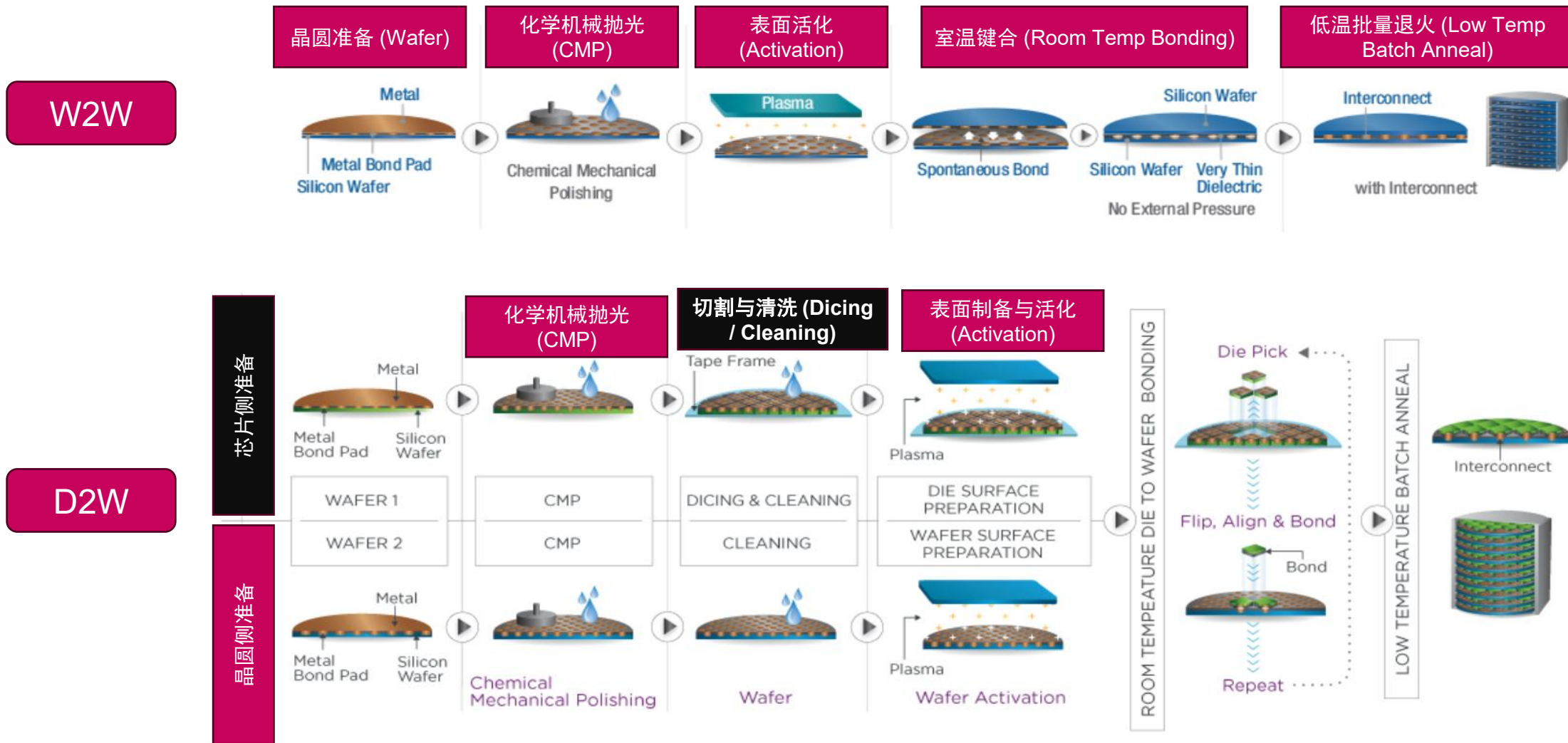
先进封装技术演进-3D

- 根据被键合对象（晶圆、裸片）的不同，衍生出多种集成架构。每种架构在吞吐量、良率成本、对准精度及应用场景上存在显著的工程权衡。
- W2W是指将两整片晶圆（通常均为300mm规格）进行整面对准、键合，随后进行减薄、TSV露头及切割的工艺流程。这是目前混合键合技术中最成熟、应用最广泛的形态：
 - **吞吐量优势：** W2W具有天然的并行处理优势。一次对准和键合操作即可同时完成成千上万个芯片的连接。相比之下，D2W需要重复数千次“拾取-放置”动作
 - **洁净度与工艺控制：** 完整的晶圆表面更平整、更易于清洗，且没有芯片切割带来的硅粉尘污染风险，因此更容易实现混合键合所需的“零颗粒”环境。
 - **良率耦合陷阱 (Yield Compounding)：** 假设Top Wafer和Bottom Wafer的良率均为90% (0.9)，在W2W过程中，位置固定的“坏芯片”可能会与“好芯片”结合，导致最终成品良率仅为0.81 (0.9×0.9)。如果堆叠层数增加到4层，良率将骤降至65% ($A = 0.9^4$)。因此，**W2W仅适用于工艺极度成熟、良率极高 (>95%甚至98%) 且上下层芯片尺寸一致的场景。**
- D2W是将经过测试、切割后的独立裸片 (KGD)，逐个拾取并以高精度键合到目标晶圆 (Target Wafer) 的特定位置上，**是实现高性能异构集成 (Heterogeneous Integration) 的方案：**
 - **已知好芯片 (KGD) 策略：** D2W的核心价值在于良率解耦。制造方可以预先测试裸片，只将合格的芯片 (KGD) 键合到目标晶圆的合格区域上。这意味着大尺寸逻辑芯片（如CPU、GPU）不会因为与坏芯片键合而报废。对于良率波动较大的先进制程节点，D2W是经济上唯一可行的方案
 - **灵活性：** D2W允许将不同尺寸、不同功能 (Logic, Memory, Analog)、不同工艺节点（如3nm Logic + 28nm I/O）、甚至不同材料 (Si + GaAs/GaN) 的芯片集成在同一个中介层或晶圆上
 - **吞吐量与成本挑战：** 逐个芯片的“拾取-清洗-活化-对准-放置”过程极其耗时。传统倒装焊机的速度 (UPH) 在混合键合的高精度要求下会大幅下降。目前，提升D2W贴片机的UPH是设备厂商的竞争焦点
 - **颗粒管理：** 晶圆切割过程 (Sawing) 会产生大量微米级硅屑。而在混合键合中，任何微粒都会导致键合失效。因此，**D2W必须配合成本较高的等离子切割 (Plasma Dicing) 或激光隐形切割技术，以及复杂的单芯片清洗工艺**



先进封装技术演进-3D

图表: W2W vs D2W



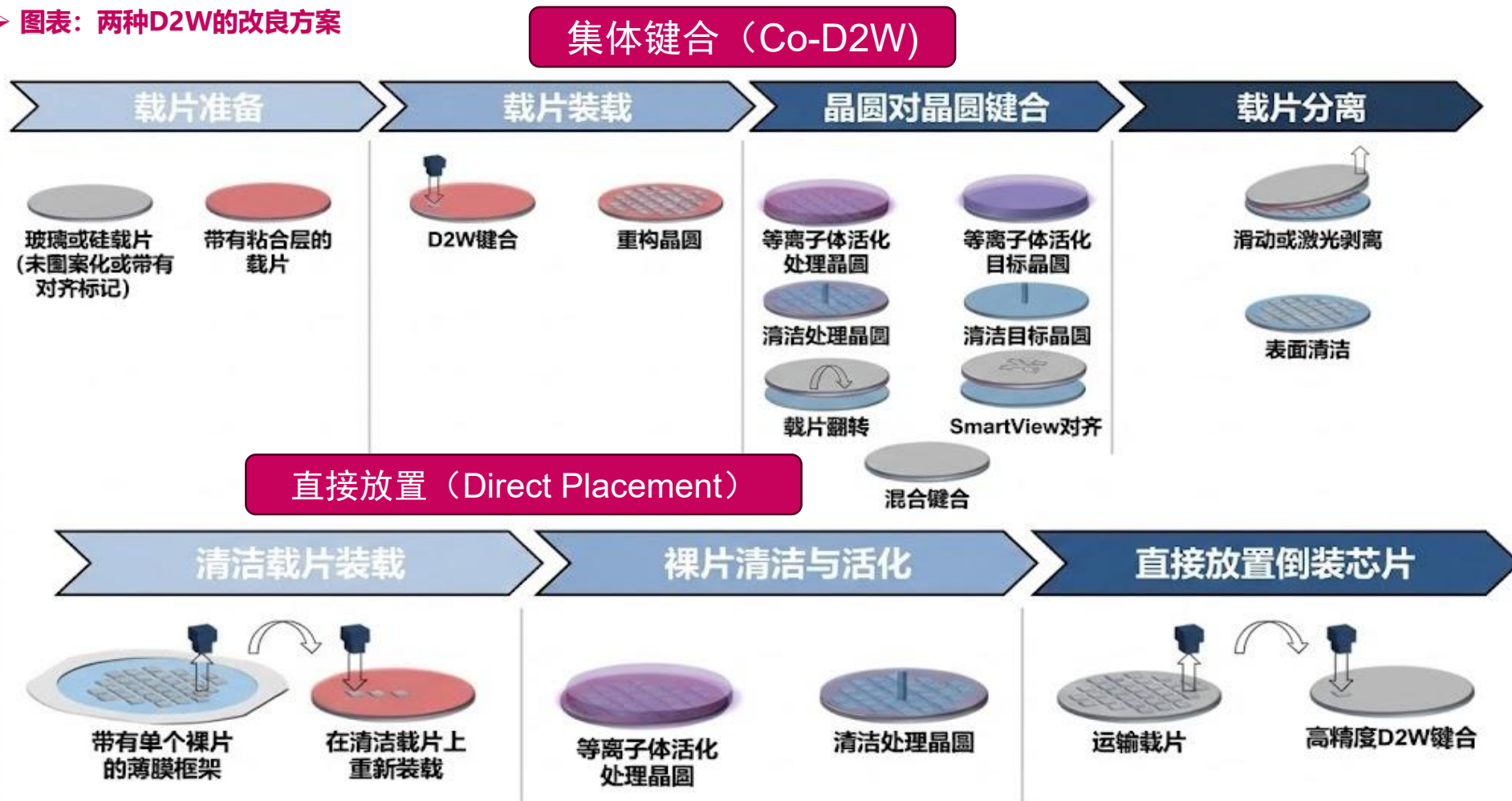
先进封装技术演进-3D

为了解决吞吐量问题及异质键合，D2W领域分化出两条技术路线，即直接放置（Direct Placement D2W）与集体键合（Collective D2W / Co-D2W）。

直接放置：使用超高精度的倒装焊机将清洗活化后的芯片直接键合到目标晶圆上。这是目前的主流方案，也是AMD 3D V-Cache（基于TSMC SoIC）采用的量产方案。其挑战在于如何在保持<200nm精度的同时提高贴片速度

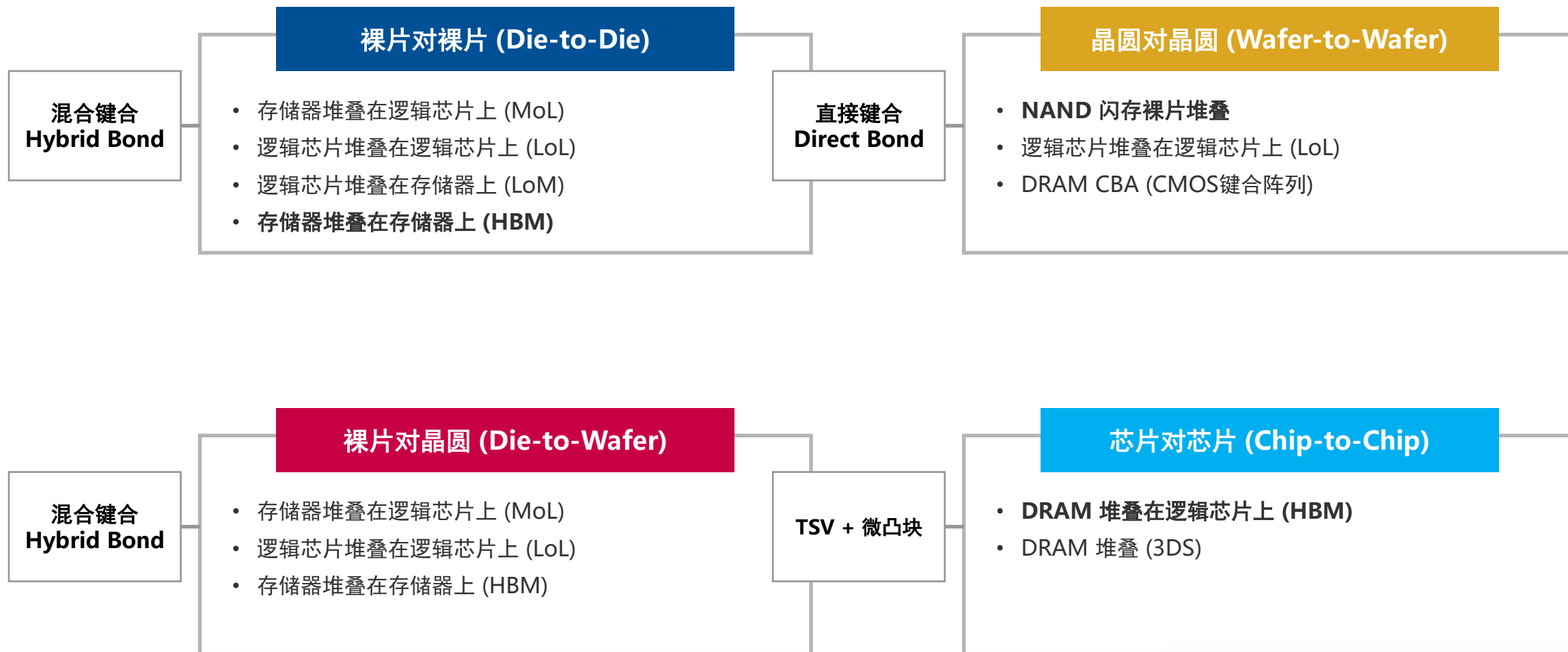
集体键合：这是一种折衷方案，旨在结合W2W的高效率和D2W的KGD优势。首先将筛选后的KGD以较低精度快速放置在一个临时载板（Carrier Wafer）上；清洗整个载板；然后将载板与目标晶圆进行一次性W2W键合；最后剥离载板。优势在于将费时的高精度对准转化为一次性的晶圆级清洗，且可以在载板上进行更彻底的晶圆级清洗。**EVG和IMEC大力推崇此路线，认为其在大规模生产中更具潜力。**

图表：两种D2W的改良方案

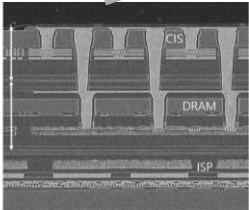
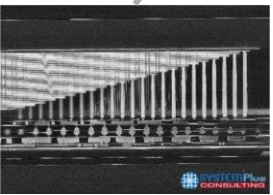

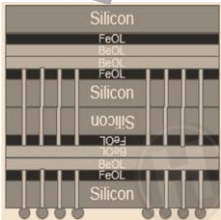


先进封装技术演进-3D

➢ 应用视角来看，TSV+微凸块是当前HBM键合的主要应用领域，其次3D NAND闪存裸片堆叠已是长江存储Xtacking工艺的核心工艺。未来，随着3D存储、MoL、LoL应用拓展，混合键合及直接键合或奖成为主流。



先进封装技术演进-3D

器件堆叠结构	背照式图像传感器	存储				Logic			MicroLED 显示器
		3D NAND Flash	HBM Stacks	DDR6+	下一代存储	SoC 分区 (SoC Partitioning)	缩放/微缩 (Scaling)		
	光电二极管 + DRAM + 逻辑芯片	NAND 模块 + 外围电路 (Periphery)	12层以上堆叠	DRAM 下方的外围电路	MRAM, FeRAM, PCM 上的外围电路	SoIC	SRAM + 逻辑芯片	背面供电网络 (BS PDN) (5nm 节点)	外延层 (EPI layer) + 逻辑芯片
键合工艺	W2W	W2W	W2W and/or D2W	W2W	W2W	W2W and/or D2W	W2W	W2W	W2W and/or D2W
	混合键合	混合键合	混合键合	直接键合	直接&混合键合	Hybrid	混合键合	直接键合	直接或者混合键合
互联间距	2μm → 1μm	2μm → 1μm	5μm → 3μm	2μm → <1μm	2μm → <1μm	9μm → 2μm	2μm	取决于光刻机精度	2μm → <1μm
成熟度	大规模量产	大规模量产	研发阶段	研发阶段	研发阶段	产能爬坡	产能爬坡	产能爬坡	产能爬坡
									
	Sony (System+)	YMTC (System+)	Xperi (ECTC2020)	IMEC (PTW21)	IMEC (PTW21)	TSMC WoW SoIC	IMEC Collaboration	IMEC Collaboration	



www.jyzq.cn



全国统一客服电话: 95372

此文件版权归金元证券股份有限公司所有, 未经许可任何单位或个人不得复制、翻印。

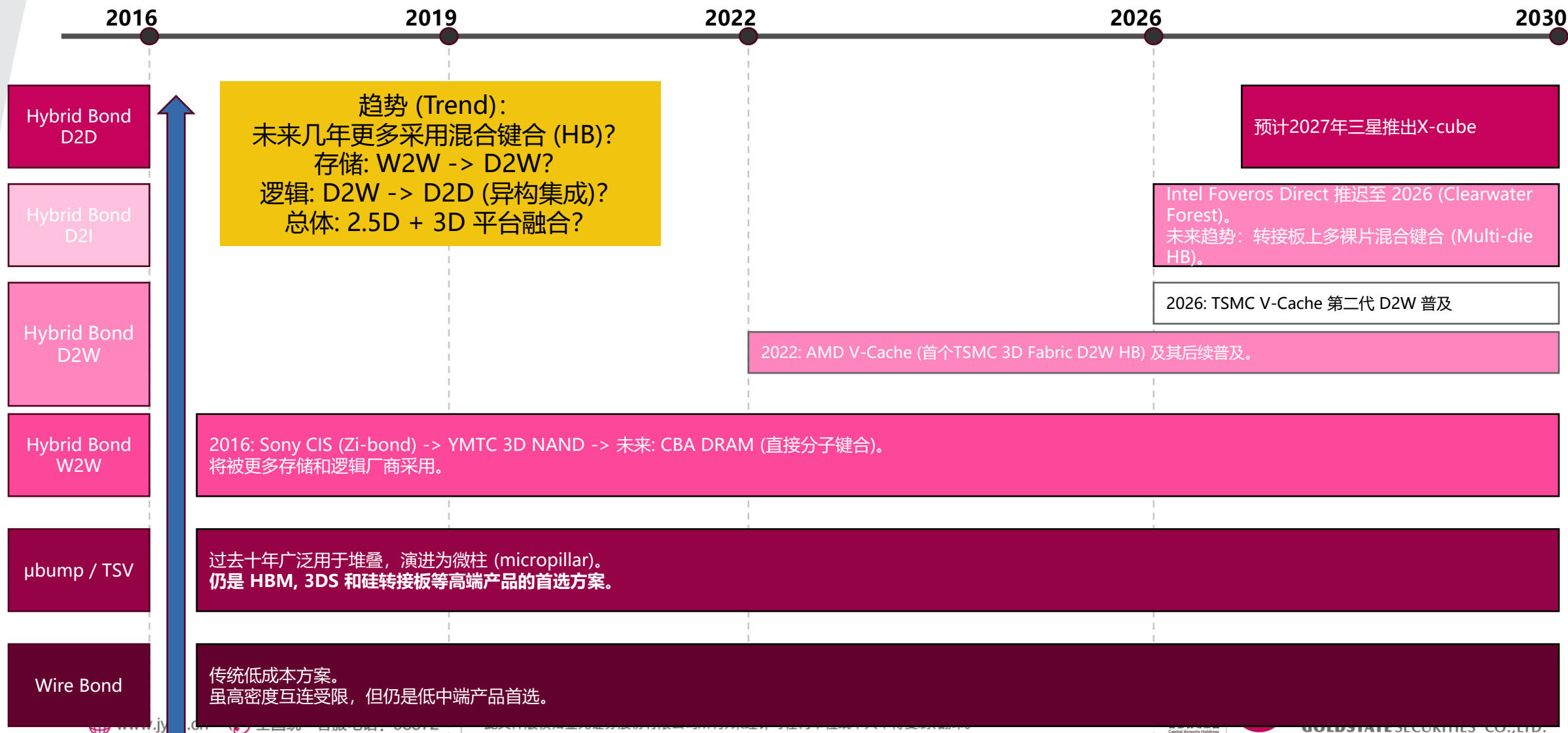


首都机场集团
Capital Airport Group



金元证券股份有限公司
GOLDSTATE SECURITIES CO., LTD.

先进封装技术演进-3D



先进封装技术演进-3D

- **3D键合工艺**依赖一系列高精度设备，其主要包括**键合机、表面活化处理设备、对准系统、清洗系统、化学机械抛光设备**。
- **键合机 (Wafer Bonder)**：负责晶圆/晶片的精密对准与键合。在W2W或D2W流程中，键合机提供纳米级对准精度、可控的压力、温度和真空/气氛环境。典型设备如EVG的GEMINI系列、SUSS MicroTec的各类键合机、拓荆科技的Dione 300系列等。
- **表面活化处理设备**：在键合前对器件表面进行活化处理，提高成键能力。**常见方法为等离子体活化或化学预处理**。等离子体活化工具（如EVG AAT、拓荆科技Propus 300等）可以在较低温度下形成高反应性表面，使得后续键合所需温度大幅降低。
- **对准系统 (Aligner)**：用于实现两侧晶圆/晶片的精准对位。W2W键合通常使用双显微镜/光学对准系统，实现亚微米级对位精度；D2W或D2D应用中，也可使用高速贴片机进行对位。
- **清洗系统**：用于键合前后去除颗粒和有机残留物。晶圆清洗设备（如Megasonic/超声波清洗、SC-1/SC-2化学槽、Deionized水冲洗系统）确保表面洁净度满足空洞容忍度。目前也出现专门的晶片清洗+活化一体化系统（如EVG 320 D2W系统）用于D2W工艺，它结合了微米级颗粒清洗与表面活化功能，提高键合良率
- **化学机械抛光设备 (CMP)**：用于使键合层达到所需的表面光洁度。该设备可将介电层表面抛光至约0.5纳米的平整度与光洁度，铜焊盘则需达到约1纳米的标准。
- **对于芯片对晶圆及芯片对芯片键合，需采用等离子切割技术以高精度、低损伤的方式将芯片从晶圆上分离**。这种非接触式工艺能保持芯片边缘的完整性，对确保芯片对晶圆和芯片对芯片混合键合的质量至关重要，因而成为该工艺的必要步骤。

先进封装技术演进-3D

化学机械抛光 (CMP)

CMP



华海清科

CMP materials



安集科技

鼎龙股份

切割 / 划片 (Dicing)

Dicing



光力科技

中微公司

北方华创

临时键合胶



清洗与表面活化

清洗与表面活化



盛美上海

北方华创

拓荆科技

对准 + 芯片到晶圆键合

Die attach & D2W bonder



拓荆科技

晶圆到晶圆 (Wafer-to-Wafer) 键合机



拓荆科技

键合后检测

量测与缺陷设备



拓荆科技

中科飞测

先进封装市场-2.5D、3D工艺价值量高

➢ 从单位封装成本来看，不含基板价值量中由于使用Si中介层、模塑中介层嵌入硅桥技术良率、工艺复杂度相对较高，单位价值量较高。存储应用中，HBM整体高于CBA DRAM、3D NAND闪存。

➢ 图表：2.5D、3D封装单位成本

封装技术名称	2024 ASP (\$)	2025 ASP (\$)	ASP/mm ² (\$/mm ²)
HBM	14.31	17.18	0.21
3D堆叠	4.85	4.85	0.06
3D NAND 闪存堆叠	2.59	2.57	0.02
CMOS 键合阵列 DRAM (CBA DRAM)	2.59	2.57	0.02
3D SoC	6.52	6.26	0.05
硅中介层 – CoWoS-S (台积电技术)	201.9	174.3	0.1
模塑中介层-CoWos-L(台积电技术)	245.88	248.33	0.13
模塑中介层 – FOEB (扇出型嵌入式桥接)	184.41	186.25	0.1
模塑中介层 – InFo-LSI (台积电技术)	33.17	39	0.06
有源硅中介层 – Foveros (英特尔技术)	29.65	29.5	0.32
嵌入式硅桥 – EMIB (英特尔技术)	3.03	3.03	0.07
模塑中的嵌入式硅桥	0.92-2.04	0.92-2.04	0.02-0.04
嵌入式硅桥 + 有源硅中介层 – Co-EMIB	41.07	40.81	0.07(EMIB)-0.32(Foveros)
超高密度扇出 – CoWoS-R (利用RDL布线, 台积电技术))	70.63	71.7	0.074
超高密度扇出 – InFO-oS (基板上扇出, 台积电技术) 及其他	27.77	27.8	0.054



先进封装市场-2.5D、3D工艺价值量高

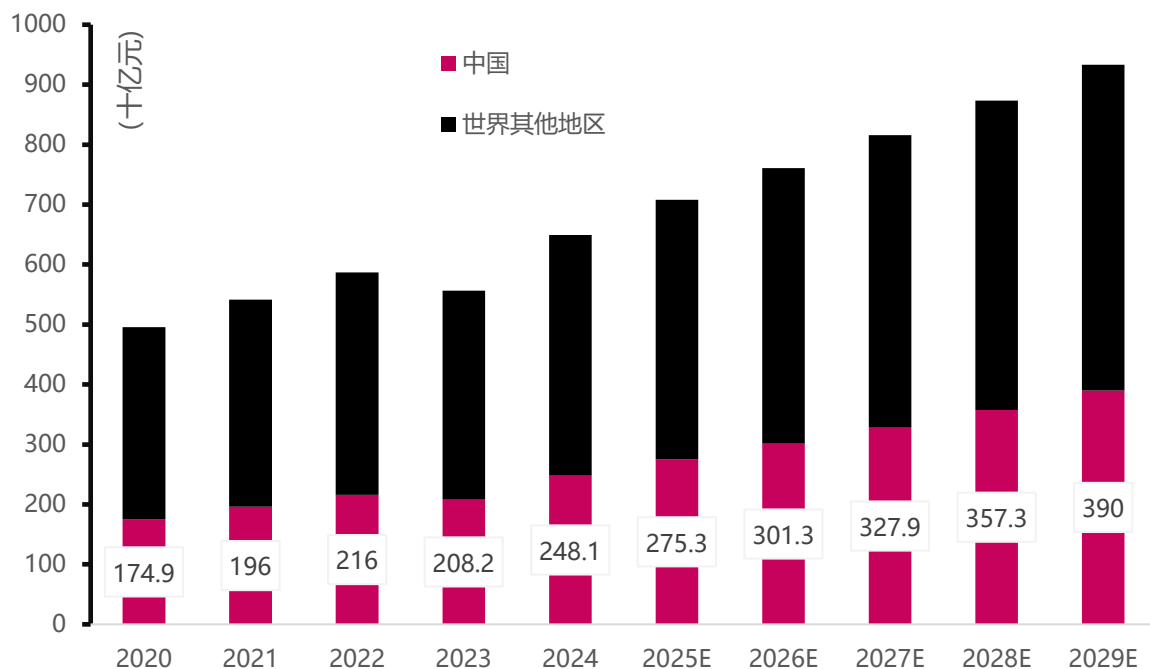
- 细分工艺来看，2.5D封装的主要成本来自于中介层制造（Si、模塑中介层及硅桥）和封装基板，而3D封装的主要成本来自键合工艺。

封装技术名称	ASP (\$)	成本构成项	成本占比 (%)
3DS	4.85	DRAM Stack TSV + Micro-Bumping	69
		Final Packaging	31
HBM	17.18	Micro-Bumping	36
		HBM Stacking	49
		HBM TSV	15
3D NAND STACK	2.57	Top Metals & Passivation	22
		Hybrid Bonding	78
MOLD INTERPOSER	39-248	Probe & Dicing Cost	4
		CoWoS-L/InFO-L/FOEB RDL Cost	44
		Dies Molding & RDL Cost	52
3D SOC	6.26	Top Metals & Passivation	17
		Hybrid Bonding	83
SI INTERPOSER	174.3	Raw Wafer Cost (Si 300mm)	6
		Interposer BEOL cost (3 ML)	12
		TSV Manufacturing Cost	4
		Micro-Bumping Cost	2
		Yield losses Cost	1
		CoW Total Cost	75
ACTIVE SI INTERPOSER	29.5	TSV, Micro-Bumping & Assembly	16
		Active Interposer FE	84
UHD FO	28-72	UHD FO Process	22
		IC Substrate + Final Packaging	78
EMBEDDED SI BRIDGE	0.92-3.03	EMIB	22
		Final Packaging	45
		IC Substrate	32
EMBEDDED SI BRIDGE + ACTIVE SI INTERPOSER	40.81	Final Packaging	32
		Active Interposer	20
		Si Bridge	8
		IC Substrate	39

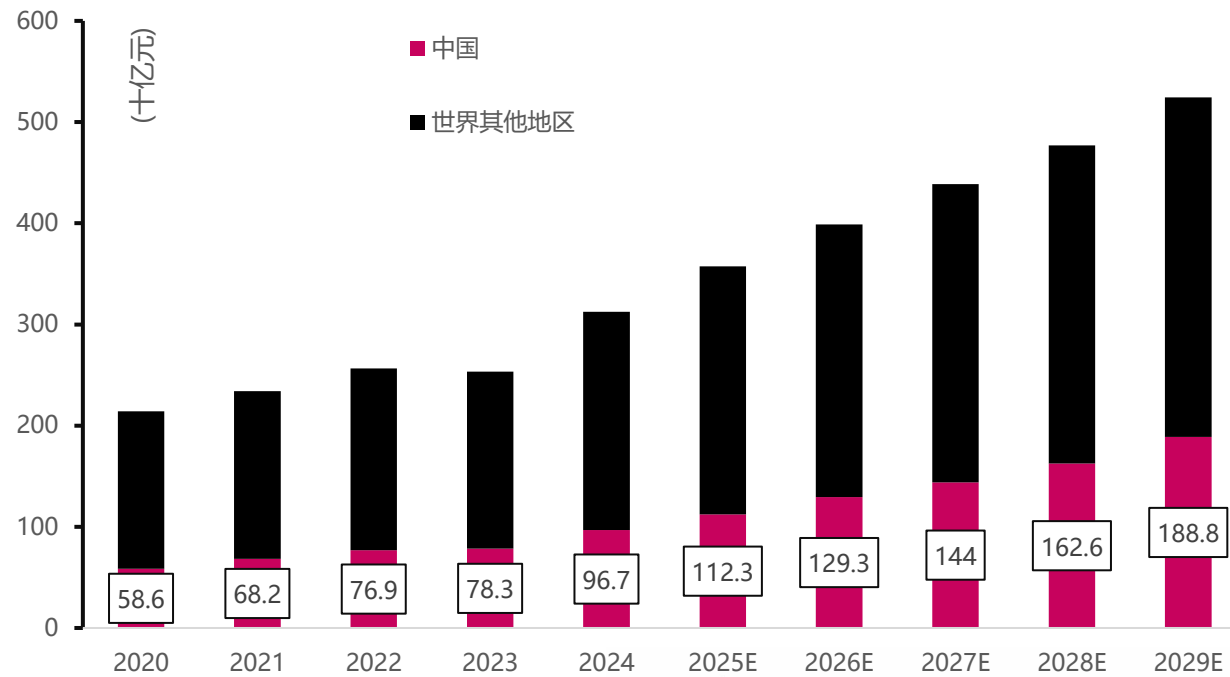
先进封装市场-2.5D、3D工艺价值量高

- 2024年，中国半导体封测市场规模约2481亿元，约占全球封装测试市场规模的38.20%。预计2029年，中国半导体封测市场达3900亿元，2024-2029年年复合增长率达9.50%，约占全球封测市场的41.80%（预计2029年，全球封测市场规模9330亿元）。
- 先进封装市场方面，2024年中国先进封装市场约967亿元，占全球市场规模的30.95%，随着AI对高性能算力芯片的需求，预计2029年中国半导体先进封装测试市场将达到1888亿元，2024-2029年年复合增速达14.30%，2029年预计中国先进封测市场将占全球市场规模的36%。

➢ 图表：全球及中国封测市场



➢ 图表：先进封装



相关公司

- 我们认为，先进封装是解决当前数据中心算力及边缘AI的核心解决方案。一方面在于成本优势，当前通过先进制程提升晶体密度的边际效益在减弱，IC的设计成本及先进制程CapEx导致性能/瓦/美元（Perf/Watt/Dollar）指标偏低。而通过小芯粒+堆叠的方式不仅能够节省成本，更重要的是随着封装体积的增大，要想继续拓展性能边界，需要在光刻机的 reticle（掩模版曝光视场）的限制下实现密度提升，小芯粒+异质集成不仅可以实现性能提升，且能够实现“混合制程”的成本优势。
- 而2.5D封装及3D封装正是为了解决芯片间互连性能及系统性能的关键技术，当前2.5D封装通过不同材料的中介层实现芯片间互连，其中模塑体嵌入硅桥方案可满足系统拓展，且相较于硅中介层有一定的成本优势。
- 但是，随着微凸点尺寸持续缩小，面临焊料桥接（Solder Bridging）风险、金属间化合物（Intermetallic Compounds, IMC）、底部填充（Underfill）工艺的挑战，键合可以实现3D晶圆堆叠、亚10微米超细间距以及无凸块设计，实现了极高的I/O密度、带宽和存储密度，从而带来显著的系统性能提升与封装高度缩减；然而，其大规模量产仍受制于多重技术瓶颈，包括对表面平坦化与洁净度的苛刻要求、互连对准的高精度门槛、高温退火可能引发的器件损伤，以及目前在吞吐量、良率控制、电气测试标准和纳米级计量检测方面存在的局限性。所以，封装设备，包括键合机、对准系统、表面活化系统、CMP设备材料或将迎来新一轮升级。
- 相关公司：
 - 1、设备厂商：拓荆科技（688072.SH）、中微公司（688012.SH）、盛美上海（688082.SH）、光力科技（300480.SZ）、北方华创（002371.SZ）、中科飞测（688361.SH）等
 - 2、材料：鼎龙股份（300054.SZ）、安集科技（688019.SH）、飞凯材料（300398.SZ）等
 - 3、OSAT：盛合晶微（未上市）、长电科技（600584.SH）、深科技（000021.SZ）等



风险提示

- 技术风险：当前2.5D封装及3D封装种类较多，分歧较大，可能导致相关公司业绩受到影响
- 供应链风险：当前封装设备及零部件国产化率不足，导致设备供应不及预期
- AI需求不及预期：国内数据中心算力需求不及预期，导致产业链需求不及预期



投资评级说明

金元证券行业投资评级标准：

增持：行业股票指数在未来6个月内超越大盘；

中性：行业股票指数在未来6个月内基本与大盘持平；

减持：行业股票指数在未来6个月内明显弱于大盘。

金元证券股票投资评级标准：

买入：股票价格在未来6个月内超越大盘15%以上；

增持：股票价格在未来6个月内相对大盘变动幅度为5%~15%；

中性：股票价格在未来6个月内相对大盘变动幅度为-5%~+5%；

减持：股票价格在未来6个月内相对大盘变动幅度为-5%~-15%；。

免责声明

本报告由金元证券股份有限公司（已具备中国证监会批复的证券投资咨询业务资格）制作。本报告所载资料的来源及观点的出处皆被金元证券认为可靠，但金元证券不保证其准确性或完整性。该等信息、意见并未考虑到获取本报告人员的具体投资目的、财务状况以及特定需求，在任何时候均不构成对任何人的个人推荐。投资者应当对本报告中的信息和意见进行独立评估，并应同时考量各自的投资目的、财务状况和特定需求，必要时就法律、商业、财务、税收等方面咨询专业财务顾问的意见。对依据或者使用本报告所造成的一切后果，金元证券及/或其关联人员均不承担任何法律责任。投资者需自主作出投资决策并自行承担投资风险，任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。

本报告所载的信息、材料或分析工具仅提供给阁下作参考用，不是也不应被视为出售、购买或认购证券或其他金融工具的要约或要约邀请。该等信息、材料及预测无需通知即可随时更改。过往的表现亦不应作为日后表现的预示和担保。在不同时期，金元证券可能会发出与本报告所载意见、评估及预测不一致的研究报告。

金元证券的销售人员、交易人员以及其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。金元证券没有将此意见及建议向报告所有接收者进行更新的义务。金元证券的自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。

在法律许可的情况下，金元证券可能会持有本报告中提及公司所发行的证券头寸并进行交易，也可能为这些公司提供或争取提供投资银行业务服务。因此，投资者应当考虑到金元证券及/或其相关人员可能存在影响本报告观点客观性的潜在利益冲突。投资者请勿将本报告视为投资或其他决定的唯一信赖依据。

本报告的版权仅为金元证券所有，未经书面许可任何机构和个人不得以任何形式转发、翻版、复制、刊登、发表或引用。

