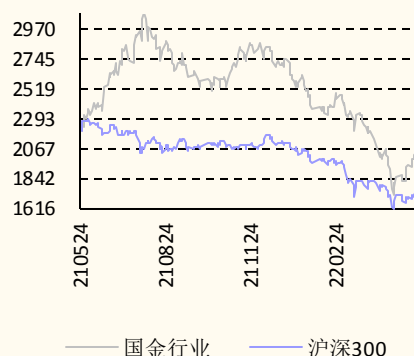


半导体 买入 (维持评级)

行业深度研究

市场数据(人民币)

市场优化平均市盈率	18.90
国金半导体指数	2030
沪深 300 指数	4078
上证指数	3147
深证成指	11455
中小板综指	11572



相关报告

- 1.《全球 WiFi 芯片的竞争格局及市场潜力-WiFi 行业深度报告》，2022.5.16
- 2.《芯片动能强，估计修正近尾声，库存风险升-《2021 年&202...》，2022.5.3
- 3.《看好新能源、智能汽车产业链机会-创新中心双周报》，2022.4.5
- 4.《乌俄关系紧张，半导体材料供应风险分析-半导体材料点评》，2022.2.13
- 5.《EDA 工具：芯片产业基础，国产快速突围-EDA 行业深度》，2021.12.31

股价提前反映半导体周期拐点的逐一浮现

投资建议

- **行业策略：**在每年机器产生庞大同比倍数增长的数据，在电动车及服务器年需 20% CAGR 增量的模拟芯片及 20% CAGR 增价的 MCU, CPU, GPU, AI, WiFi 芯片，所造成的结构性缺货下，我们估计 2023 年全球逻辑芯片行业的下行周期将相对稳定（5-7% 的同比增长），但预期各种短料芯片交期将缩短，全球/国内芯片库存月数将调整，同比营收增长将趋缓，12“成熟制程产能扩产可能将超过需求增量，虽然拐点讯号已经逐一浮现，但全球及国内半导体龙头公司估值调整已大幅领先基本面趋缓，我们维持半导体行业买入评级。
- **推荐强应用组合：**我们推荐强应用的龙头芯片公司澜起科技（服务器内存接口，PCIE Gen 4/5 retimer, 津逮 CPU），圣邦股份（模拟芯片龙头，产品组合分散），斯达半导（车用电力功率 IGBT），兆易创新（国内存储器设计龙头），及北方华创（国产替代芯片设备龙头）。

行业观点

- **全球半导体强需求的结构性缺货：**1) 从 2019 年开始，机器制造的数据每年以倍数同比增长，这些庞大数据需要透过人工智能芯片来处理分析及将有用数据收集在存储器芯片；2) 自驾电动车驱动未来数年超过 20% 的车用芯片增长 CAGR，其中包括各种模拟芯片（电源管理，马达驱动，讯号调整，收发器），电力功率（MOSFET, IGBT, SIC），MCU，边缘运算 CPU，FPGA, AI GPU, 及存储器需求；3) 元宇宙，大数据，迭代竞争加速驱动服务器主要及配套芯片需求超过 20% CAGR，其中包括 5/4/3/2nm x86/ARM CPU/AI GPU, DDR5/6, PCIE Gen 5/6 retimers, FPGA, NV Link, Infinity Fabric, CXL 及高速网络芯片等。
- **拐点讯号逐一浮现：**1) 我们预期车用及服务器用短料芯片（模拟，MCU, 电力功率，FPGA, 网络芯片）交期将从一季度的 40-50 周，逐季减少 3-5 周，到 2022 年四季度的 30 周，要是季度交期减少多于预期表示拐点明确；2) 因担心晶圆代工涨价及怕被踢出晶圆代工排队队伍的重复下单，我们预期全球逻辑芯片库存月数从 2021 年底的不到 3.5 个月，朝超过合理芯片库存月数迈进（全球 4 个月，国内芯片设计客户 5 个月合理，目前达 6.5 个月）；3) 我们预期 2022-2024 年除了车用，服务器，工业用芯片需求将稳健增长超过 20% 外，其余弱应用（手机，笔电，电视，成熟消费性电子产品）芯片营收同比增长将低于 10 个点，甚或衰退；4) 我们预期 2023 年 8“成熟制程扩产（同比增 8%）及 12”先进制程扩产（同比增 15-20%）符合需求，但 12“成熟制程扩产（同比增 16%）将超过需求（同比增 8%）近 8 个点；5) 从晶圆代工涨价是否能持续来看拐点。

风险提示

- 乌俄战争扩大的风险，国内生产链中断的风险，全球通胀升息降需求的风险，晶圆代工产能供过于求及芯片库存月数超平均的风险。

邵广雨 联系人
shaoguangyu@gjzq.com.cn

赵晋 分析师 SAC 执业编号：S1130520080004
zhaojin1@gjzq.com.cn

内容目录

一、全球半导体强需求的结构性缺货	4
1.数据驱动芯片需求	4
2.电动 / 自驾驱动车用芯片的结构性缺货	4
3.服务器芯片加速叠代更新造成的结构性缺货	7
二、半导体拐点讯号已逐一浮现	11
1.从交期变化看拐点	11
2.从全球及国内合理芯片库存月数看拐点	11
3.从同比营收增长变化看拐点	13
4.从产能扩充同比变化看拐点	14
5.从晶圆代工涨价是否能持续来看拐点	15
三、芯片行业风险	15

图表目录

图表 1: 机器数据量图表	4
图表 2: 车用芯片交期 - MCU, 模拟	5
图表 3: 车用芯片交期 - MOSFET, IGBT 电力功率	5
图表 4: 模拟芯片增量, MCU 芯片增值	5
图表 5: 全球电动车及 L3-L5 自驾车销量的占比变化	6
图表 6: 每车半导体价值及车用半导体占全球份额变化	6
图表 7: 丰田碳化硅 PCU 与硅 PCU 体积对比	7
图表 8: OBC 的硅基方案与 SiC 方案 BOM 的比较	7
图表 9: 服务器芯片 5 年增值一倍	8
图表 10: 服务器芯片短料交期周数	8
图表 11: 新世代 x86 服务器 CPU 规格比较表	9
图表 12: 英特尔 CPU + 英伟达 GPU AI 服务器架构	9
图表 13: 英伟达 Grace CPU + GPU AI 服务器架构	9
图表 14: CXL 技术用途	10
图表 15: 云端 AI 加速器比较表	10
图表 16: 桌机及笔电芯片短料交期周数	11
图表 17: 智能手机芯片短料交期周数	11
图表 18: 全球各应用逻辑芯片库存月数	12
图表 19: 国内逻辑芯片库存月数	12
图表 20: 国内芯片设计龙头一季度库存月数, 环比, 同比	12
图表 21: 全球晶圆代工各应用营收同比	13
图表 22: 全球逻辑芯片各应用营收同比	13
图表 23: 全球半导体行业链及产品应用增长同比预测	14
图表 24: 全球晶圆代工产能增长预测	14

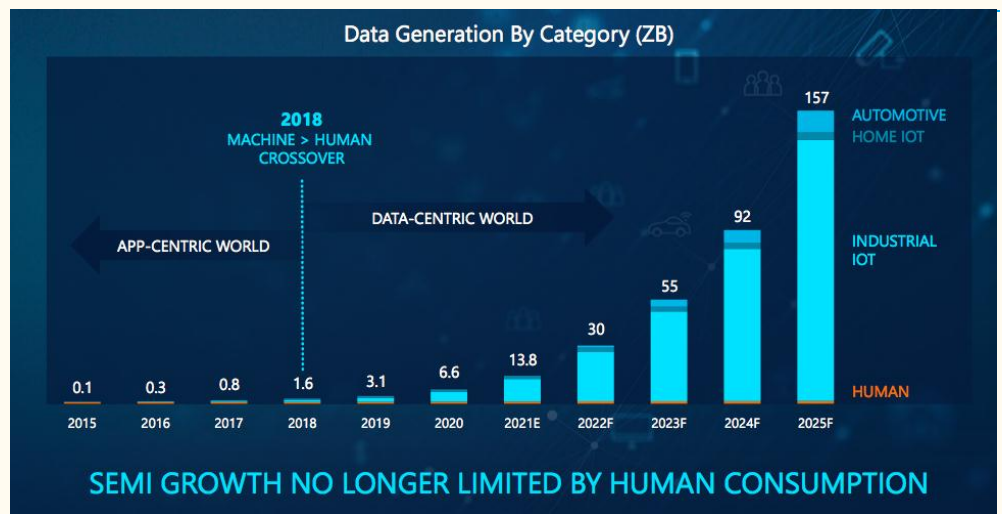
图表 25: 全球晶圆代工价格变化.....	15
------------------------	----

一、全球半导体强需求的结构性缺货

1. 数据驱动芯片需求

- 根据应用材料提供的资料，机器所产生的数据量在 2018 年首次超越人类所创造的数据量，从 2019 年，每年几乎以倍数的幅度来增加，从 2020 年到 2025 年，全球数据增量将达到 157 Zetabytes (1 Yotabyte=1000 Zetabytes; 1 Zetabyte=1000 Exabytes; 1 Exabyte=1000 Petabytes; 1 Petabyte=1000 Terabytes; 1 Terabyte=1000 Gigabytes)，5 年有 89% 复合增长率。以这样的速度增长，我们很快在 2028 年就会看到超过 1 Yotabyte 的数据增量。这么庞大的数据增量，不可能用人工来处理分析，必须运用各种具备高速运算的人工智能芯片来过滤，处理分析，训练及推理，这将持续带动 7nm 以下高速运算 HBM 存储器，3D NAND, CPU, AI GPU, FPGA, 网络芯片晶圆代工的需求，及顺势带动成熟制程的配套芯片如电源管理芯片，PCIE Gen 4/5 retimer 等的需求。

图表 1：机器数据量图表

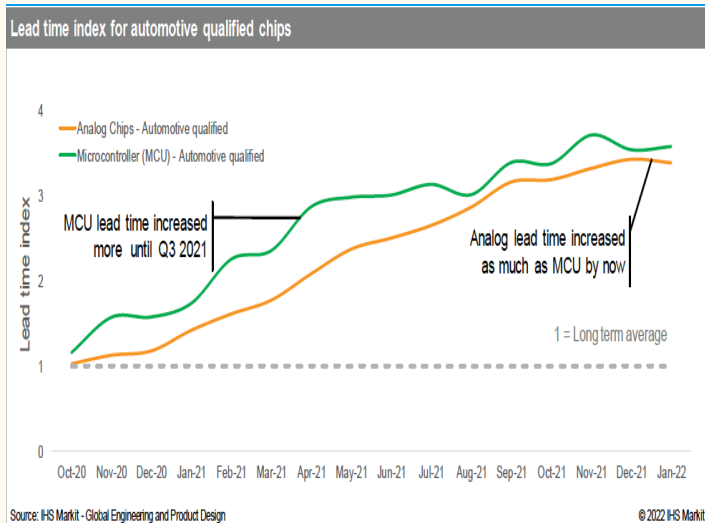


来源：应用材料，国金证券研究所

2. 电动 / 自动驾驶车用芯片的结构性缺货

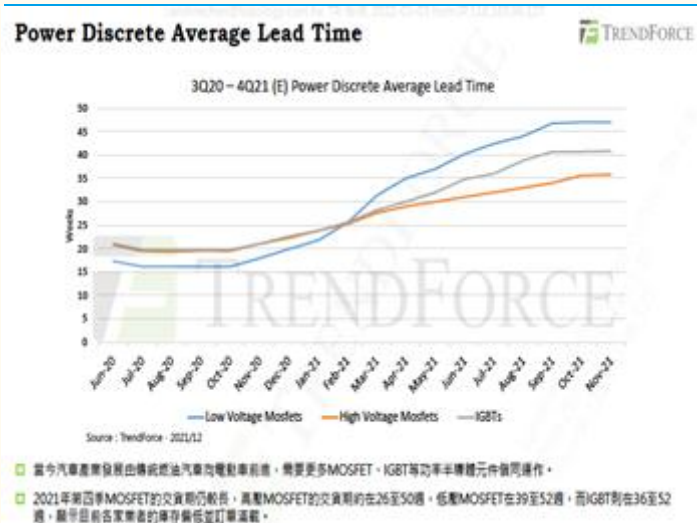
- 2021 年全球车厂最大的噩梦就是缺芯减产，在全球疫情趋缓后，因消费者为怕染病减少搭乘地铁，公交，出租车，但转而采买新车及二手车，全球市场因此从 2020 年三季度开始大幅复苏，新能源电动车及 L2 以上 ADAS 燃油车的增量又远超传统油车，再加上 NXP 及英飞凌因美国德州奥斯丁 8“厂因 2021 年初暴风雪大停电而停产，日本 Renesas 厂因 2021 年 3 月火灾而停产，而电动车及 L2 以上 ADAS 及自动驾驶系统渗透率的提升，更造成车用 MCU (每增加一台自驾感测器，雷达，激光雷达，毫米波雷达就需要一个 MCU)，电源管理芯片，电力功率等芯片缺口越来越大，因为结构性缺货，多家产业机构如 Boston Consulting Group, IHS Markit, TrendForce 的研究预测，1Q22 全球车用 MCU, 模拟，电力功率芯片的交期仍超过 40 周或三个季度，估计 2021 年全球有将近 1000 万辆车的需求因缺芯停产而无法满，这相当于贡献 2022 年超过 10 个点的增长。

图表 2：车用芯片交期-MCU，模拟



来源：IHS Markit，国金证券研究所

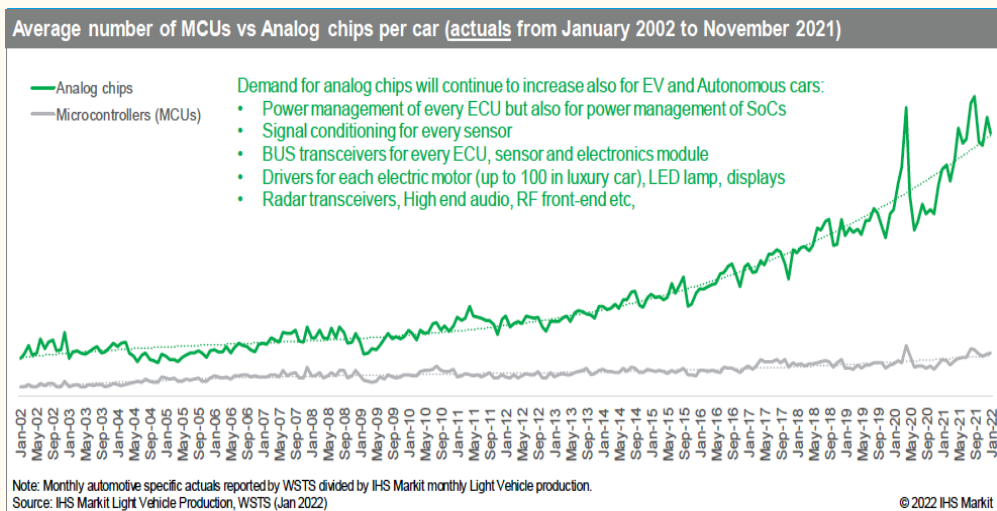
图表 3：车用芯片交期-MOSFET, IGBT 电力功率



来源：TrendForce，国金证券研究所

- 我们保守测算全球车厂营收在 2022 / 2023 年仍有 9% / 10% 的增长，但全球车用半导体市场在 2022/2023 年却有 24% / 20% 的同比增长。这对 2022 年全球半导体增长有 2-3 个点的贡献。每车芯片价值在 2022/2023 年有近 10-15% 的增长，而每车半导体价值因电动车（估计 2035 年有 70% 渗透率）及 ADAS/自动驾驶（估计 2035 年有超过 30% 渗透率）的渗透率增加而提升，因先进制程芯片成本增加而提升，因芯片数量增加而提升（主要指模拟芯片），我们认为未来 15 年，全球车用半导体增长将明显高于车厂营收增长平均达 10-15 个点，全球车用半导体市场于 2020-2035 年复合成长率可能超过 20%（1-3% CAGR 来自于全球车市出货量的成长，9-11% CAGR 来自于每车车用芯片数目增长，8-10% CAGR 来自于芯片平均单价提升）。我们认为未来 15 年重点关注的强应用公司包括英伟达 Orin 的出租车自动驾驶芯片系统，Tesla 的乘用车 FSD/D1 自动驾驶芯片系统，高通，地平线，黑芝麻的自动驾驶芯片，Lumentum 的激光雷达光源，WolfSpeed, Rohm 及 STMicroelectronics 的 6"/8" 碳化硅解决方案。

图表 4：模拟芯片增量，MCU 芯片增值

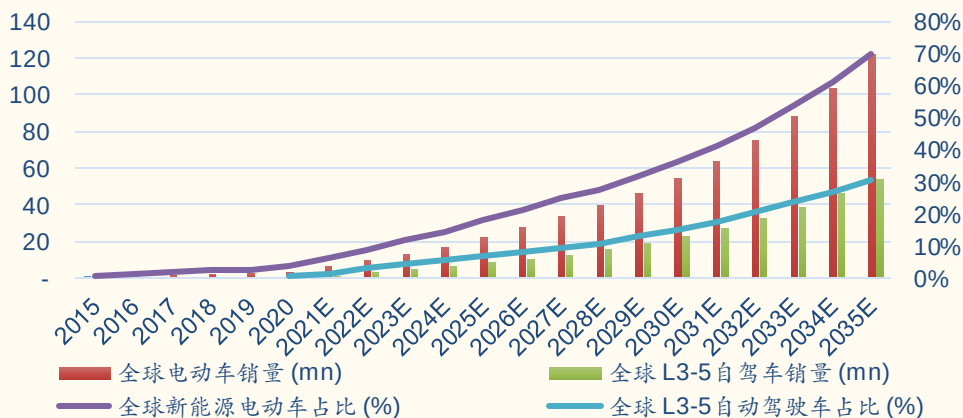


来源：IHS Markit 3/8/2022, WSTS Jan 2022，国金证券研究所

- 人驾到自驾，重点在成本及视觉 / AI 芯片技术：很多产业专家说未来的自动驾驶就像装了四个轮子的智能手机，以自动驾驶的难度及半导体配置而言，我们不同意这说法，我们认为自动驾驶像是装了四个轮子的智能 AI 服务器（如果透过远端控制软件来协作，自动驾驶更像装了四个轮子的智能集群系统），Gartner 在 2019 年四季度预测在 2023 年，全球有近 74.6 万自驾

车，而目前使用激光雷达来作为视觉功能的 SAE L4-L5 的自驾车成本要超过 10 万美元，昂贵的激光雷达感测元件价格 5 万美元以上，所以很难普及到自用车，像是特斯拉不使用激光雷达，但透过 3 颗前置摄像头（60, 150, 250 公尺视觉距离），1 颗后置摄像头（50 公尺视觉距离），4 颗前后侧边摄像头（80-100 公尺视觉距离），12 颗环绕车身的超音波感测器（感测距离 8 公尺），及一颗前置雷达（160 公尺视觉距离）推出的 L3 等级 FSD 自驾解决方案，整体额外自驾功能成本应该不超过 2 万美元。我们估计于 2035 年全球超过 30% 的汽车销量将具备 L3-L5 的自动驾驶功能，未来 15 年的复合增长率达到 30-35%。

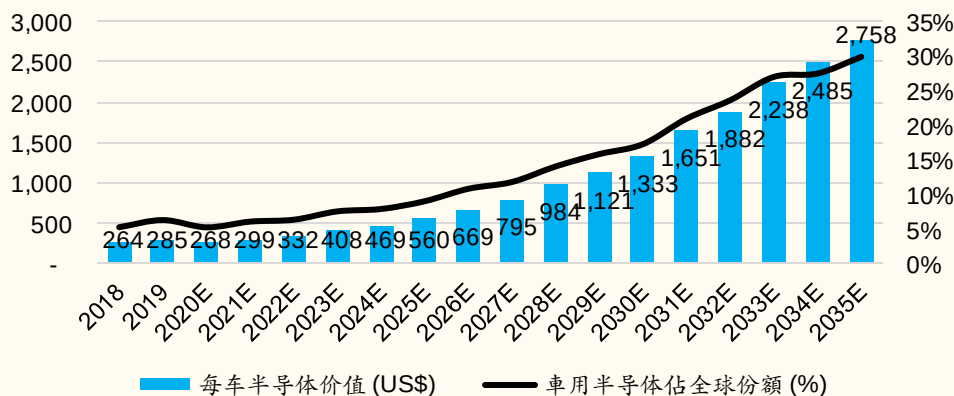
图表 5：全球电动车及 L3-L5 自驾车销量的占比变化



来源：The Boston Consulting Group (BCG)，国金证券研究所

- 我们认为汽油引擎车转马达电动车，接着是由人驾转 SAE 3-5 级自驾车的占比提升，加上电动车及自驾车的技术演进（耗能降低，电池密度提升，电源转换系统重量降低，摄像头，感测器，雷达，激光雷达数量提升，及人工智能芯片运算能力提升但要求耗能持续降低），这些技术演进将逐步拉升每台电动车及自驾车的半导体价值，这两大驱动力对全球车用半导体公司及产业未来二十年将产生重大影响，我们先前估计全球车用半导体市场于 2020-2035 年复合成长率应有机会超过 20%（主要系增加 AI GPU, FPGA, ASIC，激光雷达，以太网络，MCU，模拟芯片，IGBT，碳化硅，电源管理芯片的价值及数量），远超过全球半导体市场在同时间的复合成长率的 7-9%，约占全球半导体市场的份额将在 2035 年达到 30%（从 2021' s 不到 10 个点），每车半导体价值从 2020 年的 268 美元，暴增 10 倍到 2035 年的 2,758 美元。

图表 6：每车半导体价值及车用半导体占全球份额变化



来源：TrendForce，国金证券研究所测算

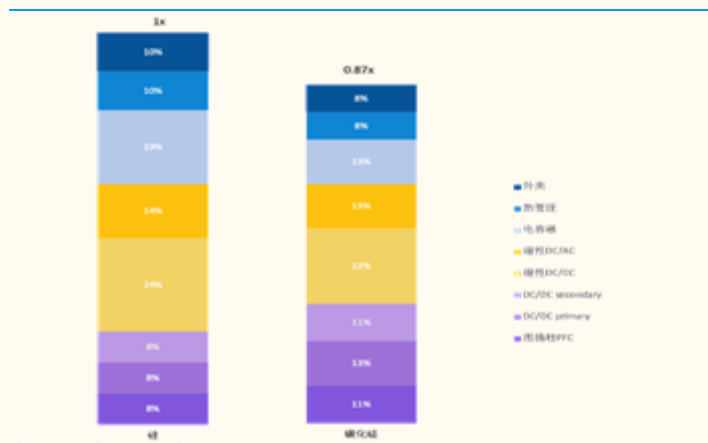
- 电动车 SiC 碳化硅方案带来五大优势：目前电动车（不包括 48V MHEV）系统架构中涉及到功率器件的组件包括：电机驱动系统中的牵引逆变器（Traction Inverter, DC - AC, 直流转交流电）、车载充电系统（OBC, On-board charger）、800V 高功率电源转换系统（车载 DC-DC 转换器）和非车载充电桩。电动汽车采用碳化硅解决方案可以带来五大优势：1. 可以提高开关频率降低能耗。采用全碳化硅方案逆变器开关损耗下降 80%，整车能耗降低 5%-10%；2. 可以缩小动力系统整体模块尺寸，以丰田开发的碳化硅 PCU 为例，其体积仅为传统硅 PCU 的五分之一 3. 在相同续航情况下，使用更小电池，减少无源器件使用，降低整体物料成本。以电动汽车的 6.6kW 双向 OBC 为例，典型 DC-AC 部分包括四个 650V IGBT、几个二极管和一个 700- μ H 电感，占材料清单成本的 70% 以上。通过使用四个 650V SiC MOSFET 实现，只需要 230 μ H 的电感。这比基于 IGBT 的设计降低了将近 13% 的材料清单成本。4. 缩短电池充电时间，由于更高的充电功率和更小的电池，可以大幅缩短电动车充电时间。5. 在 600 度 C 的工作温度下有高度稳定的晶体结构，击穿场强是 IGBT 的 10 倍多，导通损耗小，2.5 倍于硅材料的热导系数。

图表 7：丰田碳化硅 PCU 与硅 PCU 体积对比



来源：丰田，国金证券研究所

图表 8：OBC 的硅基方案与 SiC 方案 BOM 的比较



来源：wolfspeed，国金证券研究所

3. 服务器芯片加速叠代更新造成的结构性缺货

- 我们估计全球服务器厂商营收在 2022 / 2023 年有 5% / 7% 的增长，但全球服务器半导体市场在 2022 / 2023 年却有 20% / 25% 的增长。这对 2022/2023 年半导体增长各有 4-6 个点的贡献。每台服务器芯片价值在 2022 年有超过 12% 的增长，主要系 AI 智能服务器比重的提升对 AI GPU 需求也有提升，而 Intel 7 及 AMD 5nm 的 CPU 的芯片面积大增估计也对成本及价格进一步提升，PCIe Gen 5 retimer, 及 DDR5, DDR5 内存接口芯片的采用，都对每台服务器芯片有增量，增价效果。因为 Intel, AMD, Nvidia 在云端，边缘运算，企业，政府，运营商端的技术叠代竞争加速，我们认为未来 10 年，全球服务器半导体增长将明显高于服务器厂商营收增长平均达 10-15 个点，全球服务器半导体市场于 2021-2035 年复合成长率达 20%（4-5% CAGR 来自于全球服务器数量成长，2-3% CAGR 来自于每台服务器芯片数目增长，12-14% CAGR 来自于芯片平均单价提升）。根据美国半导体设备龙头应用材料的预估，从 2020 到 2025 年，全球每台服务器的半导体芯片价值将增加一倍到 5600 美元，相当于 20% 的 5 年复合增长率。归因于每台服务器的芯片数量增加，芯片面积变大，良率变差，处理层数增加，交货周期拉长所造成的结构性缺货，TrendForce 的研究预测，2022 年 3 月全球服务器用 FPGA, 电源管理芯片，高速网络芯片，电力功率芯片的交期仍超过 40 周或三个季度。

图表 9：服务器芯片 5 年增值一倍

Semi content per unit	2015	2020	2025F
HIGH END SMARTPHONE	\$100	\$170	\$275
AUTO (GLOBAL AVERAGE)	\$310	\$460	\$690
DATA CENTER SERVER (CPU + ACCELERATOR)	\$1,620	\$2,810	\$5,600
SMARTHOME (GLOBAL AVERAGE)	\$2	\$4	\$9

SILICON CONTENT GROWING AS EVERYTHING GETS SMARTER

来源：应用材料，国金证券研究所

图表 10：服务器芯片短料交期周数

Server (Ranking of Components by Severity of Shortage)			
Rank	Item	Company	Lead Time (Unit: Week)
1	FPGA	Lattice	around 50
2	PMIC	Texas Instruments, ADI (Maxim), Microchip	40~50
3	LAN chip	Broadcom	40
4	MOSFET	Onsemi	40

来源：TrendForce，国金证券研究所

- Intel 7 Sapphire Rapids: 英特尔于今年三月份小量出货的新一代服务器芯片 Intel 7 CPU Sapphire Rapids, 虽然是第一次改用 Chiplet 架构, 但在整合四颗 372mm² 芯片 (每颗芯片 (tile) 有 14-20 CPU 核心) 面积的 chiplets 后, 整体 56-80 核心芯片面积估计高达 1488mm², 是上个世代 Ice Lake 产品的三倍大左右, 所以定价应该会比 Ice Lake 贵三倍左右, 我们以为因为核心数不足, 芯片面积过大, 良率较差, 成本高, ABF 大载板缺货 (层数变多, 面积变大, 良率变差) 等因素, 英特尔 2022 年要靠着 Sapphire Rapids 收回过去两年失去的给 AMD 份额, 我们觉得还是有难度。那为何英特尔不改其 CPU 设计为小芯片架构, 我们研究认为主要系英特尔的架构已经有 2, 4, 8 颗服务器 CPU Socket 插槽的主机版, 如果每颗 chiplet 的 CPU 有 8 颗芯片 (Tile), 那 8 个 CPU 的服务器主机板就会有 64 颗芯片 (Tile) 要互相连结, 那 CPU 跟 CPU, 芯片跟芯片的沟通都要透过 UPI 通道 (Ultra Path Interconnect, 最高传输速度可达 16GT/s), 那 UPI 肯定会成为 CPU 芯片之间沟通的瓶颈。还有就是, 英特尔整合了两个 DDR5 的控制芯片到一颗芯片 (Tile) 中, 4 颗芯片就是 8 个存储器通道, 如果是 8 颗芯片, 每颗芯片太小就无法容纳两个 DDR5 的存储器通道控制芯片, 这样也会有瓶颈。而 AMD 是把存储器控制芯片放在中心的 I/O 控制芯片上, 所以每颗 CPU 多加几颗芯片没有关系。而且, AMD 的服务器架构主要用 1, 2 颗服务器 CPU Socket 插槽的主板, 如果每颗 chiplet 的 CPU 有 8 颗芯片 (Tile), 那 2 个 CPU 的服务器主机版就仅有 16 颗芯片要互相连结, 沟通上就简单许多。
- TSMC 5nm Genoa: 即使是 AMD 超威从 7nm 服务器芯片 Milan 微缩到 5nm Genoa, 因为微缩有限又加入更多晶体管, 面积从 8 芯 64 核 592mm² CPU 核心芯片面积大幅增加 40% 到 12 芯 96 核 828mm² CPU 核心芯片面积, 但每颗 8 核心芯片面积只有从 74mm² 微缩 7% 到 69mm²。所以虽然 I/O 控制芯片从 GlobalFoundries 的 14nm 416mm² 面积, 微缩到台积电 6nm 的 263mm², 但整体芯片面积 (加总 CPU 核心芯片及 I/O 控制芯片) 还是增加了 8 个点, ABF 大载板面积也因为需要更多的接脚数 (Socket 脚数从 7nm LGA 4094 到 5nm LGA 6096), 至少增加了 20% 的面积。我们估计今年四季度 Sapphire Rapids 及 Genoa 两个芯片加起来有机会占全球 x86 服务器 CPU 出货量超过 15%, 单价提升轻易超过 10 个点。

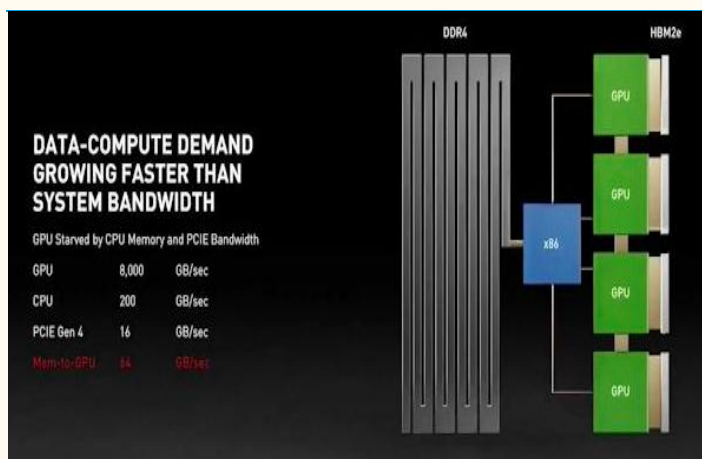
图表 11: 新世代 x86 服务器 CPU 规格比较表

	Intel			AMD		
Product name	Ice Lake	Sapphire Rapids	Emerald Rapids	Milan	Genoa	Bergamo
平台	Whitley	Eagle Stream	Eagle Stream	Zen 3	Zen 4	Zen 4
制程节点	10nm+	Intel 7nm	Intel 7	7nm EUV	5nm EUV	5nm EUV
核心数	32-40 cores	56 cores	64 cores	8x8 64 cores	12x8 96 cores	16x8 128 cores
线程数	64-80 Threads	112 Threads	128 Threads	128 Threads	192 Threads	256 Threads
核心芯片面积	550-600mm ²	372mm ² x 4 chiplets: 1488mm ²	372mm ² x 4 chiplets: 1488mm ²	74mm ² x8 chiplets: 592mm ²	69mm ² x 12chiplets: 828mm ²	
I/O 芯片				GF 14nm 416mm ²	TSMC 6nm 263mm ²	
量产时点	2Q21	2Q22	2Q23	2Q21	2Q22	2Q23
存储器通道数	8x DDR 4- 3.2Ghz	8x DDR 5	8x DDR 5	8x DDR 4- 3.2Ghz	12x DDR 5- 5.2Ghz	
DRAM 密度/CPU	128GBx8x2=2TB	128GBx8x2=2TB	128GBx8x2=2TB	128GBx8x2=2TB	128GBx12x2=3TB	
PCIe Gen 4/5	64x PCIe G4	80x PCIe G5	80x PCIe G5	128x PCIe 4	128x PCIe 5	
DDR	DDR 4 3200MT/s	DDR 5 4800MT/s	DDR 5/5600MT/s	DDR 4 3200MT/s	DDR 5 5200MT/s	
TDP	250-270W	270-350W		280W	400W	
Socket	LGA 4189	LGA 4677	LGA 4677	LGA 4094	LGA 6096	

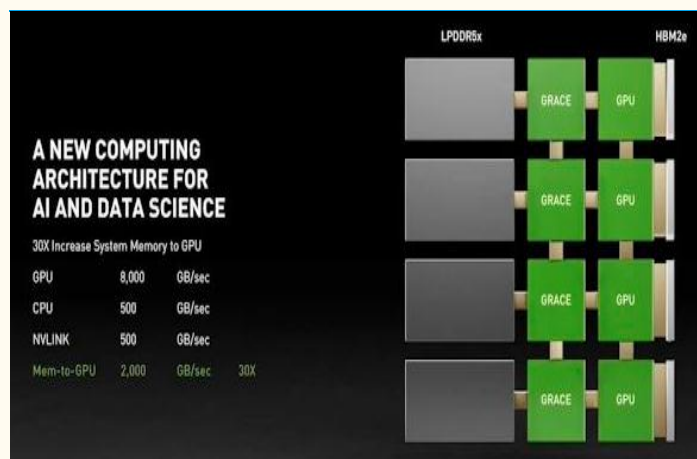
来源: 各公司官网, 国金证券研究所

- AI 服务器拉升英伟达云端 AI 芯片需求: 这几年因为 Nvidia 英伟达的 GPU 大量被使用在人工智能的云端辨识系统, 从每台服务器加两片高速运算 GPU 卡, 4 片到 8 片 GPU 卡都有, 让 Nvidia 在数据中心芯片市场的份额从 2018-2019 年的不到 10%, 到 2021 年的超过 20%, 全球人工智能服务器占比也逐年提升至 2021 年近 10%, 但光就 AI 推理及训练加速器而言, 英伟达应该有超过 95% 以上的份额。英伟达的 4nm 芯片 H100, 芯片面积虽然 812mm², 最大耗电量达 700W, 但其在浮点半精度, 单精度, 双精度稀疏及理论峰值运算都明显优于同业, 为了让 H100 的 DGX 人工智能服务器系统发挥效能, 还要另外配备非常昂贵的 320-640GB HBM (高频宽内存) 给 AI GPU 使用, 为了控制使用价格昂贵的 HBM 让总成本不要失控, 英伟达决定在 2023 年初推自家的设计的 ARM CPU (Grace) 配合其高速 NVlink 通讯网络 (500GB/秒传输速率), 可以让 CPU/GPU 共同分享 DDR4/5/HBM 存储器达到快取一致性(Cache Coherency), 我们认为未来哪家公司能率先推出 AI 服务器达到 Cache Coherency, 这就是未来 AI 服务器竞争者的决胜点之一。

图表 12: 英特尔 CPU + 英伟达 GPU AI 服务器架构



图表 13: 英伟达 Grace CPU+ GPU AI 服务器架构

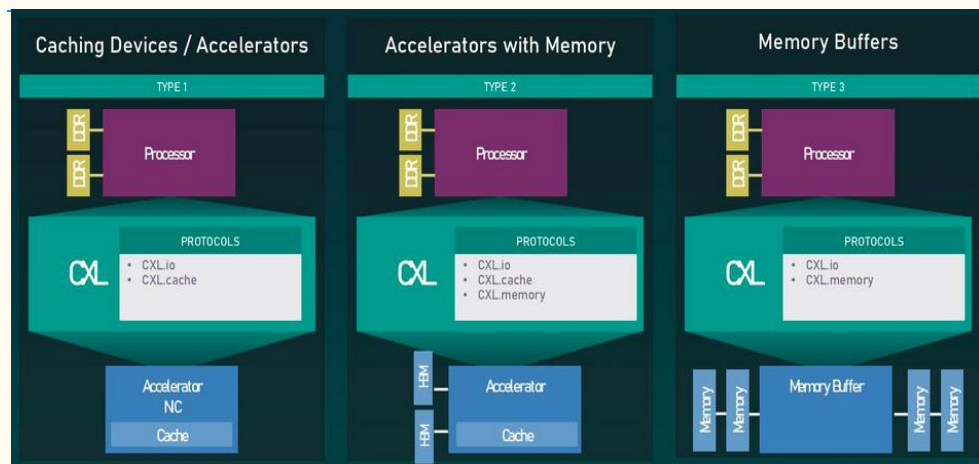


来源: Nvidia, 国金证券研究所

来源: Nvidia, 国金证券研究所

- 英特尔要自推 AI GPU 抢份额: 英特尔今年要推出的 AI GPU 加速器 Ponte Vecchio (1000 亿晶体管, 47 芯片, tsmc 5nm x16, 8 核心=128 核心 (compute tile), intel 7nm x2 640mm2 (base tile), tsmc 7nm x2 (link)) x4 再搭配 2 颗 Sapphire Rapids CPU 的人工智能服务器系统, 也需要庞大的 DDR5 DRAM 及 HBM DRAM。不同于 Nvidia 的 NVlink, AMD 的 Infinity Fabric 这两者都是独家技术封闭架构, 无法适用于 CPU 及 GPU 以外其他芯片的连结, Intel 支持的 CXL 可消除 CPU 与设备、CPU 与存储器之间的传输瓶颈, 创建一个可以支持 AI, smart I/O、smart NIC, CPU, 存储器模組的接口, 以服务下一代的数据中心。可惜目前 CXL 是建立在 PCIE Gen 5.0 的实体及电子层的架构上, 目前 PCIE Gen 5/CXL 最高只能提供 64GB/s 的双向传输速率, 跟 NVlink 的 500GB/s, Infinity Fabric 的 400GB/s 有 6-8 倍的差距, 可能要等到 PCIE Gen 6/CXL 的 128GB/s 技术问世才能进一步缩短与 Nvidia 及 AMD 之间的差距。但从 2019 年 3 月 13 日开始, 英特尔携手阿里巴巴、思科、戴尔, EMC、IBM, 脸书 Meta、谷歌、惠普、华为以及微软宣布成立 Compute Express Link(CXL) 开放合作联盟, 而目前从 NVlink, Infinity Fabric retimer 的设计进度来看, 要能够领先英特尔达到 Cache Coherency 似乎还是有很大的难度。

图表 14: CXL 技术用途



来源: CXL consortium 2020, 国金证券研究所

图表 15: 云端 AI 加速器比较表

	H100 GPU	D1 ASIC	Ponte Vecchio	MI200
供应商	英伟达	Tesla	Intel	超威
应用	云端推理训练	云端推理训练	云端推理训练	云端推理训练
CPU/GPU 连结标准	Nvlink 500GB/s		PCIE Gen 6.0/CXL 256GB/s	Infinity Fabric 3.0, 400GB/s
制程工艺	TSMC 800 亿晶体管 4nm CoWoS, 812mm2	500 亿晶体管, 7nm, 645mm2	1000 亿晶体管, 47 芯片, tsmc 5nm x16, 8 核心=128 核心 (compute tile), intel 7nm x2 640mm2 (base tile), tsmc 7nm x2 (link)	2 x CDNA2 架构, TSMC 6nm, 2x290=580 亿晶体管,
浮点 FP16 半精度 Matrix	2000 TFLOPS	362 TFLOPS		383 TFLOPS
浮点 FP32 单精度 Vector	1000 TFLOPS	22.6 TFLOPS	45 TFLOPS	47.9 TFLOPS
浮点 FP32 单精度 Matrix	N/A			95.7 TFLOPS
浮点 FP64 双精度 Vector	N/A			47.9 TFLOPS
浮点 FP64 双精度 Matrix	60 TFLOPS			95.7 TFLOPS
最大耗电量	700W	400W		560 W

来源: 各公司官网, 国金证券研究所

- 服务器用 DDR5 即将问世: 在英特尔已于三月份量产出货支援 4800 MT/s

DDR5 的 Intel 7 服务器 CPU Sapphire Rapids, 而 AMD 即将推出的 EUV 5nm 服务器 Zen 4 CPU Genoa, 也将支援 5200MT/s 的 DDR5, 我们认为 DDR5 将比 DDR4 芯片面积及价格提升 25-30%, 意思就是消耗掉更多的内存 DRAM 芯片产能。除此之外, 为了能够梳理 CPU 与 DDR5 内存之间大量的数据存取, 整体 DDR5 模组中 DDR5 1+10 内存接口芯片比重应该会提升超过 10%, 内存接口芯片面积也会加大, 还要推新串行检测, 温度传感, 电源管理芯片等配套芯片。Trendforce 研究机构还预期 DDR5 模组的电源管理芯片因产能短缺, 可能面临缺货的窘境。我们估计 2022 年澜起, Renesas/IDT, Rambus 将分食 40%/40%/20% 的 DDR5 内存接口芯片份额, 而目前仅澜起及 Renesas/IDT 能提供三个完整的配套芯片解决方案。

二、半导体拐点讯号已逐一浮现

- 虽然目前车用, 服务器, 工业用芯片需求仍然短缺, 但 TV, 笔电, 智能手机(摄像头, 显示驱动芯片)需求不振但芯片迭代更新持续, 全球半导体需求仍然强劲, 但交期, 库存月数, 营收同比, 产能扩充的变化都会影响价格变动, 这都可能让行业出现拐点讯号。

1. 从交期变化看拐点

- 根据 HIS Markit 及 Trendforce 2022 年一季度的芯片交期报告, 目前服务器及车用模拟, MCU 芯片交期超过 40 周, 但归因于 TV, 笔电, 智能手机从 2022 年一季度以来明显转弱, 环比下跌近 15 个点, 预期 TV, 笔电, 智能手机的短料芯片(如 Type C IC, 4G SoC, A+G 感测器, OLED 显示驱动芯片)交期将陆续回归到正常水位的 10-15 周, 之前被这些应用占用的产能将逐步释放出来, 我们认为服务器及车用模拟芯片短料交期也将从 2022 年二季度开始逐季降低, 预计在 2022 年底前后下滑到 30 周左右, 明年有机会达到正常水位的 10-15 周, 一旦各种强应用及弱应用芯片产品交期都达到正常水位, 我们相信芯片及晶圆代工价格可能出现变化, 但就目前需求而言, 我们仍然无法判断明年这些强应用如服务器, 车用, 及工业用短料芯片交期会达到正常水位的 10-15 周。

图表 16: 桌机及笔电芯片短料交期周数

PC & Notebook (Ranking of Components by Severity of Shortage)			
Rank	Item	Company	Lead Time (Unit: Week)
1	Type C IC	Texas Instruments, NXP, Infineon, Parade	20~25
2	WiFi	Intel, Broadcom, Realtek	10~12
3	PMIC	Texas Instruments, MPS, ON semi, Taiwanese brands	10~20

来源: TrendForce, 国金证券研究所

图表 17: 智能手机芯片短料交期周数

Smartphone (Ranking of Components by Severity of Shortage)			
Rank	Item	Company	Lead Time (Unit: Week)
1	4G SoC	MediaTek, Qualcomm, Unisoc	30~40
2	OLED DDIC/Touch IC	Samsung (DDIC), Goodix (Touch IC), FocalTech (Touch IC)	20~22
3	A+G Sensor	BOSCH, InvenSense, STMicroelectronics	32~36

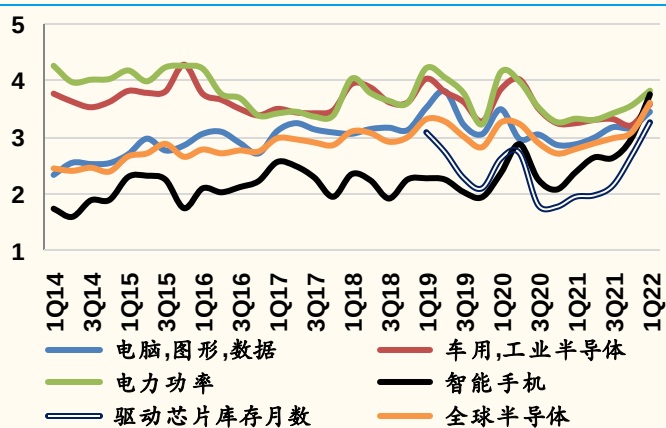
来源: TrendForce, 国金证券研究所

2. 从全球及国内合理芯片库存月数看拐点

- 从国金独家全球逻辑芯片 1Q22 数据档来看, 库存月数明显飙升, 尤其比较高的是国内芯片设计公司近 6.5 个月的库存(环比增 22%, 同比增 74%)及全球 NOR/SLC NAND/mask ROM 行业近 5 个月的库存(环比增 19%, 同比增 27%), 其他芯片行业库存月数虽然未超过合理库存的 4 个月, 但整体库存月数达 3.7 个月, 环比增长 17%, 同比增长 29%。归因于二季度上海地区的封控, 多少造成很多组装大厂的停产或降载生产, 我们预测国内芯片设计库存将超过 7.5 个月, 全球芯片设计及 IDM 库存将超过 4 个月, 我们因此预测从 2022 年三季度开始或最迟于 2023 年, 全球芯片库存月数整理即将启动。

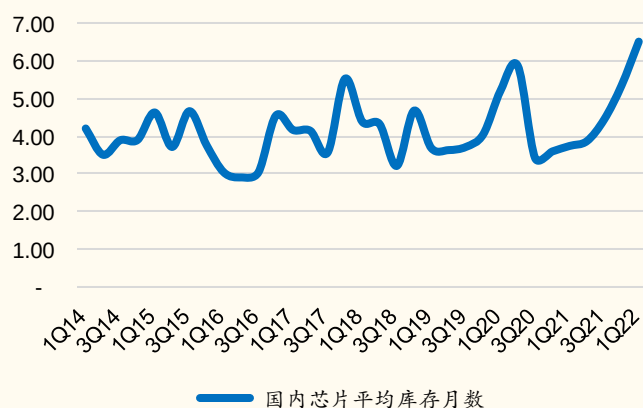
■ 以 2021 年美国白宫的产业链报告建议全球半导体供应商应保持比过去更高的库存水位，以应付各种如火灾，水灾，地震，暴风雪，战争，停电，疫情封城，地缘政治，半导体缺制程材料所造成的半导体产线中断，所以我们测算全球系统组装大厂将把合理半导体芯片库存月数从过去的 2.9-3.0 个月，到以后的 4.0 个月，我们认为国内半导体设计公司甚至将策略性地提高合理库存从 2021 年的 4 个月到以后的 5 个月。但当国内手机感测芯片大厂韦尔股份，TWS 耳机主控芯片大厂恒玄科技，手机低噪射频芯片大厂卓胜微，GPU 芯片大厂景嘉微，FPGA 芯片大厂复旦微，固态存储控制芯片大厂国科微因担心晶圆代工持续涨价，担心取消订单后要重新排队下单，担心美国技术管控加大力道，因而重复下单中高阶通用芯片库存，而纷纷公布超过 7 个月的芯片库存，同比增加近倍后，及国内主要芯片设计公司，公布 6.5 个月的加权平均历史新高芯片库存月数，环比增加 22%，同比增加 74%。我们认为当晶圆代工 2023 年涨价后继无力时，这些高库存重点客户也将成为 8“/12”成熟制程晶圆代工的第一波芯片砍单客户。

图表 18：全球各应用逻辑芯片库存月数



来源：各公司公告，国金证券研究所

图表 19：国内逻辑芯片库存月数



来源：wind，国金证券研究所

图表 20：国内芯片设计龙头一季度库存月数，环比，同比

公司	行业	1Q22 库存	环比%	同比%
澜起科技	DDR5 内存接口芯片	2.4	10%	-58%
思瑞浦	模拟芯片	2.4	-4%	-41%
中颖电子	MCU、OLED 驱动	2.6	-19%	25%
力芯微	消费类电源管理芯片	2.8	-17%	37%
创耀科技	接入网网络芯片	3.2	229%	N/A
芯朋微	家电、工业用电源管理芯片	3.7	47%	22%
扬杰科技	功率	3.7	-8%	43%
富瀚微	IPC、ISP 芯片	4.0	-18%	-4%
圣邦股份	模拟芯片	4.0	1%	-9%
希荻微	电源管理芯片	4.5	-26%	N/A
晶晨股份	智能机顶盒芯片	4.6	21%	146%
兆易创新	存储器，MCU	4.6	8%	76%
上海贝岭	高精度 ADC	4.8	33%	38%
芯海科技	MCU，健康监测 AIOT	5.1	24%	-25%
纳芯微	隔离芯片	5.2	N/A	N/A
艾为电子	模拟、射频芯片	5.8	50%	90%
瑞芯微	IPC 芯片	5.8	68%	146%
晶丰明源	照明电源管理驱动芯片	6.7	52%	238%

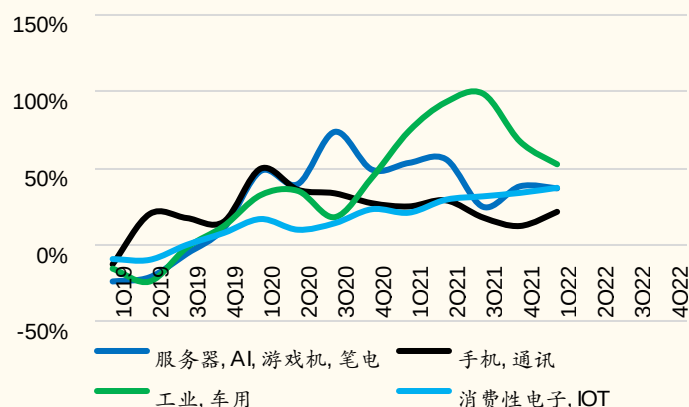
乐鑫科技	wifi 芯片	7.0	82%	83%
国民技术	MCU	7.4	63%	-25%
卓胜微	射频芯片	7.4	-17%	74%
汇顶科技	指纹识别, 电容触控芯片	8.1	127%	222%
紫光国微	智能安全, 存储, FPGA	8.2	44%	21%
韦尔股份	手机/汽车 CIS, TDDI	8.8	22%	120%
景嘉微	GPU	9.1	-13%	-24%
国科微	固态存储控制芯片, 视频监控	10.2	3%	388%
复旦微电	FPGA	10.8	15%	32%
恒玄科技	TWS 耳机主控芯片	12.6	160%	183%
寒武纪-U	智能集群系统	38.3	618%	130%
国内平均		6.5	22%	74%

来源: wind, 国金证券研究所

3.从同比营收增长变化看拐点

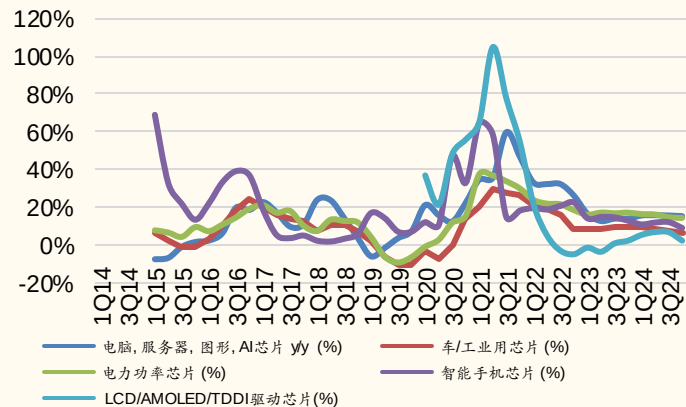
■ 我们认为半导体下行周期有分两大类, 一种是像 1997 年亚洲金融风暴, 2001 年 Y2K 后电脑更新停滞, 2008 年全球衍生性商品金融风暴的大规模的需求衰退。另一种是像 2012, 2015, 2019 年的需求增长减缓, 配合半导体产能陆续开出而造成的供过于求及芯片价格下跌调整。虽然新冠疫情持续扩散影响需求, 但疫苗接种率提高让重病比率明显下降, 全球疫情舒缓可期; 虽然通胀严重到让各国政府央行接连升息, 但半导体厂尽力将设备涨价, 折旧, 材料成本提高转嫁给需求更多产能的强应用客户; 加上摩尔定律微缩大幅趋缓, 芯片面积越来越大, 良率越来越差, 我们预期未来 10 年全球芯片制造成本及晶圆代工平均单价将逐年攀升 4-5 个点或以上, 我们将不再像过去 20 年因摩尔定律微缩加速而看到明显的芯片成本及平均芯片价格下降。我们因此测算未来 10 年 (2021-2031) 全球逻辑芯片营收将有 10-11 个点的复合增长率 CAGR, 高于过去 10 年复合增长率的 5.9%, 2023 年全球逻辑芯片营收增长虽然会比 2021 年的 24%, 2022 年的 10-15% 同比增长趋缓, 我们测算应仍有 5-7% 的营收同比增长, 而不至于步入全球逻辑芯片营收大幅衰退的需求反转周期 (1997, 2001, 2008) 以及供过于求 / 芯片价格下跌周期 (2012, 2015, 2019)。但就个别应用而言, 我们预估 LCD/AMOLED/TDDI 智能手机 / 笔电 / TV 荧幕显示驱动及触控整合芯片, TV 控制芯片, 消费型笔电芯片于 2022-2023 年期间应该会出现几个季度的同比营收衰退, 但电动 / 自驾车用及服务器用芯片营收仍会有 10 个点以上的增长。

图表 21: 全球晶圆代工各应用营收同比



来源: 各公司公告, 国金证券研究所

图表 22: 全球逻辑芯片各应用营收同比



来源: 彭博社, 国金证券研究所

图表 23：全球半导体行业链及产品应用增长同比预测

同比 y/y	2020	2021	2022	2023	2024	2025
半导体行业链						
全球半导体 Global semi sales	11%	26%	12%	11%	16%	11%
全球逻辑芯片 logic IC sales	14%	25%	14%	6%	7%	10%
全球存储器 memory IC sales	11%	31%	8%	15%	30%	15%
全球晶圆代工营收 foundry	33%	29%	28%	12%	15%	11%
产品应用面						
全球智能手机出货量 Global smartphone	-7%	7%	4%	0%	3%	2%
全球 5G 手机出货量 5G smartphone	2150%	133%	24%	38%	28%	13%
全球手机半导体营收 Smartphone IC sales	15%	31%	19%	8%	7%	6%
全球电脑/笔电出货量 Global PC	20%	12%	-15%	0%	0%	5%
全球电视出货量 Global TV	0%	0%	-10%	2%	0%	0%
全球车商营收 Automotive	-10%	11%	5%	10%	7%	5%
全球车用半导体营收 Automotive semis	-3%	39%	24%	20%	22%	15%
全球电动车出货量	20%	111%	50%	37%	31%	29%
全球服务器制造商营收 server	7%	19%	7%	5%	10%	10%
全球服务器半导体营收 server IC	26%	20%	20%	25%	20%	18%

来源：各公司公告，国金证券研究所

4.从产能扩充同比变化看拐点

- 根据我们对 2023 年全球逻辑芯片及晶圆代工营收同比增长 8% / 12% 的预估，再把单价提升 4-5 个点的因素移除，我们估计全球逻辑芯片及晶圆代工出货量同比增长 4-8%，与 8“扩产的 8% 产能同比增长同步（供需平衡），但比 12“成熟制程（14nm, 28nm, 40nm, 65nm, 90nm）扩产的 16% 产能同比增长明显偏低近 8-12 个点，2023 年可能造成供过于求，这对于中芯国际，Globalfoundries, 联电这些 12“成熟制程大厂比较负面。至于 12“先进制程，因为台积电有将近超过 90% 的晶圆代工份额，我们认为其接单生产的供需平衡模式不变，供需相对稳定。

图表 24：全球晶圆代工产能增长预测

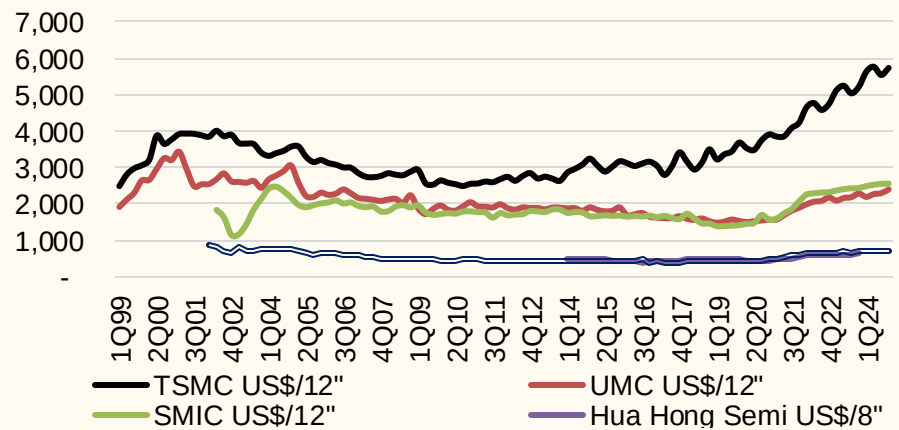
8" 供给同比	4Q19	4Q20	4Q21	4Q22	4Q23	4Q24	4Q25
TSMC	0%	4%	0%	7%	6%	6%	6%
UMC	0%	7%	3%	5%	5%	4%	3%
Vanguard	1%	17%	4%	24%	10%	6%	5%
PSMC 力积电	4%	13%	33%	0%	0%	0%	0%
SMIC	8%	7%	19%	24%	17%	12%	11%
华虹	2%	0%	0%	2%	5%	5%	5%
Global	2%	7%	6%	12%	8%	6%	6%
12“成熟供给同比							
TSMC	0%	3%	3%	3%	10%	6%	5%
UMC	9%	22%	7%	10%	15%	12%	10%
PSMC 力积电	1%	0%	0%	0%	10%	17%	15%
SMIC	4%	14%	31%	50%	24%	19%	16%
华虹		440%	344%	50%	33%	25%	20%
合肥晶合	100%	50%	133%	57%	36%	20%	17%
Global	5%	8%	13%	13%	16%	12%	10%

来源：各公司公告，国金证券研究所

5.从晶圆代工涨价是否能持续来看拐点

- 虽然 2023 年仍是充满变数的一年，但根据我们 2021 年年中策略报首次提出因摩尔定律趋缓，先进制程工艺设备成本加速增值，成熟二手制程设备短缺，我们认为未来 20 年，半导体行业将进入通胀趋势，晶圆代工年度涨价 5-10%将成为常态，所以未来 20 年的半导体行业市场的复合增长率，将比过去 20 年多个 5-10 个。
- 而台积电为反映人工，电力，芯片制造材料，设备成本明显上扬，于 2022 年 5 月 10 日对 2023 年客户正式发出产能分配名单及近 5-8%的全制程工艺节点的涨价通知数；在同一周，彭博社报道三星与代工客户正在谈 2022 年下半年 15-20%涨价；而镁光财务长最近宣布与 10 家客户签署实验性存储器芯片供应量固定，价格固定三年长约。三大半导体龙头的这些芯片价格调整都代表芯片行业的强需求仍在，有定锚的效应，2023 年半导体需求不至于崩盘，但如果其他中小型晶圆代工厂于 2023 年无法跟进涨价，代表需求拐点逐一浮现。

图表 25：全球晶圆代工价格变化



来源：各公司公告，国金证券研究所

三、芯片行业风险

- **乌俄战争的风险：**不光半导体气体供应受到影响，俄罗斯消费性电子产品需求如智能手机及电脑都明显下滑而影响全球半导体需求。
- **疫情封城管制：**大上海地区因疫情扩大而管制，对半导体厂商如中芯国际，华虹，华力，台积电上海松江 8 “，环旭生产影响不大，但物流管制，通关，下游客户停产将造成二季度半导体库存明显增加。
- **全球升息抗通胀：**尤其是美国升息，造成美元升值，对采购美国半导体设备的成本将明显提高，全球通胀也将减少对消费性电子产品的购买力，升息也会降低对半导体板块投资的吸引力。
- **供过于求的风险：**根据我们的预测 2023 年全球 12 “成熟制程产能将增加 15-20%，明显高于需求增长的 5-10 %，我们因此预测 12 ”成熟制程产能利用率将无法维持在 100%，尤其是 40/28nm 制程将有供过于求的风险。

公司投资评级的说明：

买入：预期未来 6－12 个月内上涨幅度在 15%以上；
增持：预期未来 6－12 个月内上涨幅度在 5%－15%；
中性：预期未来 6－12 个月内变动幅度在 -5%－5%；
减持：预期未来 6－12 个月内下跌幅度在 5%以上。

行业投资评级的说明：

买入：预期未来 3－6 个月内该行业上涨幅度超过大盘在 15%以上；
增持：预期未来 3－6 个月内该行业上涨幅度超过大盘在 5%－15%；
中性：预期未来 3－6 个月内该行业变动幅度相对大盘在 -5%－5%；
减持：预期未来 3－6 个月内该行业下跌幅度超过大盘在 5%以上。

特别声明:

国金证券股份有限公司经中国证券监督管理委员会批准,已具备证券投资咨询业务资格。

本报告版权归“国金证券股份有限公司”(以下简称“国金证券”)所有,未经事先书面授权,任何机构和个人均不得以任何方式对本报告的任何部分制作任何形式的复制、转发、转载、引用、修改、仿制、刊发,或以任何侵犯本公司版权的其他方式使用。经过书面授权的引用、刊发,需注明出处为“国金证券股份有限公司”,且不得对本报告进行任何有悖原意的删节和修改。

本报告的产生基于国金证券及其研究人员认为可信的公开资料或实地调研资料,但国金证券及其研究人员对这些信息的准确性和完整性不作任何保证,对由于该等问题产生的一切责任,国金证券不作出任何担保。且本报告中的资料、意见、预测均反映报告初次公开发布时的判断,在不作事先通知的情况下,可能会随时调整。

本报告中的信息、意见等均仅供参考,不作为或被视为出售及购买证券或其他投资标的邀请或要约。客户应当考虑到国金证券存在可能影响本报告客观性的利益冲突,而不应视本报告为作出投资决策的唯一因素。证券研究报告是用于服务具备专业知识的投资者和投资顾问的专业产品,使用时必须经专业人士进行解读。国金证券建议获取报告人员应考虑本报告的任何意见或建议是否符合其特定状况,以及(若有必要)咨询独立投资顾问。报告本身、报告中的信息或所表达意见也不构成投资、法律、会计或税务的最终操作建议,国金证券不就报告中的内容对最终操作建议做出任何担保,在任何时候均不构成对任何人的个人推荐。

在法律允许的情况下,国金证券的关联机构可能会持有报告中涉及的公司所发行的证券并进行交易,并可能为这些公司正在提供或争取提供多种金融服务。

本报告反映编写分析员的不同设想、见解及分析方法,故本报告所载观点可能与其他类似研究报告的观点及市场实际情况不一致,且收件人亦不会因为收到本报告而成为国金证券的客户。

根据《证券期货投资者适当性管理办法》,本报告仅供国金证券股份有限公司客户中风险评级高于 C3 级(含 C3 级)的投资者使用;非国金证券 C3 级以上(含 C3 级)的投资者擅自使用国金证券研究报告进行投资,遭受任何损失,国金证券不承担相关法律责任。

此报告仅限于中国大陆使用。

上海

电话: 021-60753903

传真: 021-61038200

邮箱: researchsh@gjzq.com.cn

邮编: 201204

地址: 上海浦东新区芳甸路 1088 号

紫竹国际大厦 7 楼

北京

电话: 010-66216979

传真: 010-66216793

邮箱: researchbj@gjzq.com.cn

邮编: 100053

地址: 中国北京西城区长椿街 3 号 4 层

深圳

电话: 0755-83831378

传真: 0755-83830558

邮箱: researchsz@gjzq.com.cn

邮编: 518000

地址: 中国深圳市福田区中心四路 1-1 号

嘉里建设广场 T3-2402