



高带宽内存争夺战开启

本报记者 张心怡 实习记者 夏冬阳

近日,SK海力士CEO宣布其2024年和2025年高带宽内存(HBM)的产能均已售罄,并预测未来专用于人工智能的“超高速、大容量、低电力”存储器需求将会激增。无独有偶,美光科技也在第二财季财报中披露,该公司2024年的HBM产能已经售罄,2025年的绝大多数产能已经分配完毕。

HBM已成为存储厂商的兵家必争之地。据Mordor Intelligence预测,从2024年到2029年,HBM市场规模预计将从约25.2亿美元激增至79.5亿美元,预测期内年复合增长率高达25.86%。为抢占这一增量市场,三星、SK海力士、美光等企业正在为HBM调配更多产能,并推动HBM的规格迭代。

头部企业积极扩大HBM产能

AI是当前半导体市场的一抹亮色。随着云服务厂商积极部署AI大模型,AI服务器进入增长轨道。TrendForce集邦咨询预估,至2026年,AI服务器(包含搭载GPU、FP-GA、ASIC等)出货量将占整体服务器的15%,成为撬动服务器市场的新增长点。

AI服务器对芯片内存容量和传输带宽提出更高要求,也拉动了高端存储芯片和高密度存储模组的出货量,尤其是HBM技术的迭代升级和应用。据市场调查机构Yole数据,今年以来HBM的平均售价是普通DRAM的5倍,加上近期SK海力士和美光的HBM产能售罄情况,进一步证实了市场对这一技术的高度认可和迫切需求。

新一轮技术竞赛已展开

自2014年首款硅通孔HBM产品问世至今,HBM技术已从HBM(第一代)、HBM2(第二代)、HBM2E(第三代)、HBM3(第四代)发展至HBM3E(第五代),最新的HBM3E是HBM3的扩展版本。

HBM代际升级主要体现在数据速率和容量密度上,HBM3的扩展版本HBM3E提供高达8Gbps的传输速度,比HBM3提高了50%;内存容量达16GB,是HBM3的两倍,且功耗更低,可满足下一代高性能计算应用的需求。

自今年2月美光先宣布量产8层24GB HBM3E并供货英伟达

在这样的趋势下,SK海力士、三星、美光正在积极扩张HBM产能,争夺市场份额。

作为HBM技术先行者的SK海力士,在2023年第四季度率先迎来盈利,其中HBM和DDR5是较大业绩支撑。财报显示,SK海力士2024年第一季度DRAM业务收入占比达61%,平均销售价格(ASP)环比上涨超过20%。SK海力士表示:“凭借HBM等面向AI的存储器技术领导力,公司提升了面向AI服务器的产品销量,同时持续实施以盈利为主的经营策略,从而实现了营业利润环比增长734%的业绩。”此外,顺应面向AI的存储器需求增长这一趋势,SK海力士决定加大于今年3月率先开始生产的HBM3E产品供

H200后,英伟达在GTC大会上确认了三代的HBM产品正处于验证阶段,而作为英伟达HBM内存长期供应商的SK海力士也宣布其HBM3E产品将从3月底开始量产。预计三星和SK海力士将围绕搭载在英伟达明年销售的下一代AI芯片中的HBM3E展开激烈的供应竞争。

在角逐HBM3E的同时,三大存储原厂已经开始布局下一代存储产品HBM4。

日前,SK海力士已宣布研发HBM4,并与台积电就HBM4研发和下一代封装技术合作签署谅解备

忘录,计划从2026年开始批量生产HBM第六代产品——HBM4。据

悉,SK海力士将采用台积电的先进工艺来提高HBM4的性能,针对搭载于HBM封装内最底层的基础裸片进行优化。台积电表示,将采用N5和N12FFC+工艺制造基础裸片,为HBM4提供更强性能和能源效率。

在HBM4研发上,三星和美光也不会轻易罢手。尽管SK海力士目前稳坐HBM市场的头把交椅,但三星正雄心勃勃地追赶SK海力士。今年以来,三星已成立两个全新HBM团队,并将HBM工作小组

转为芯片部门下的一个常设办公室,以弥补其在2019年解散HBM业务和团队的市场误判。时下,三星与SK海力士已在HBM4研发上“短兵相接”,三星HBM团队计划于今年下半年量产HBM3E,并于2025年生产HBM4,2026年实现量产。

作为HBM市场后起之秀的美光也在积极布局HBM4,据了解,美光计划在2026年至2027年间,推出容量在36GB到48GB、12/16层垂直堆叠的HBM4产品,并于2028年推出带宽增加至2TB/s以上的HBM4E。

第一季度全球IC销售额同比增长22%

将激增21%。IC库存水平在2024年第一季度趋于稳定,预计本季度将有所改善。

晶圆厂产能持续增加,预估每季将超过4000万片晶圆(以300mm晶圆当量计算),2024年第一季度产能增长1.2%,预计2024年第二季度增长1.4%。中国仍然是所有地区中产能增长率最高的国家。晶圆厂利用率,尤其是成熟节点的利用率,预计2024年上半年几乎没有复苏迹象。由于严格的供应控制,2024年第一季度memory利用率低于预期。

与晶圆厂利用率趋势一致,半导

体资本支出依然保守。在2023年第四季度同比下降17%后,资本支出在2024年第一季度继续回落11%,第二季度实现0.7%的预期增长。第二季度的这一趋势的积极转变,其中预计与内存相关的资本支出将增长8%,增长略强于非内存领域。

SEMI首席分析师Clark Tseng表示:“一些半导体领域的需求正在复苏,但复苏步伐不均衡。人工智能芯片和高带宽存储器是目前需求最高的设备,这导致了这些领域的投资和产能扩张。然而,由于人工智能芯片依赖少数关键供应商,其对IC出货量增长的影响仍然有限。”

AI服务器对芯片内存容量和传输带宽提出更高要求,也拉动了高端存储芯片和高密度存储模组的出货量。

收的71%。此外,美光预计2024财年资本支出在75亿美元至80亿美元之间,均高于去年的资本支出和此前规划,主要是为了支持HBM3E的产量增长。美光首席执行官Sanjay Mehrotra对外透露:“我们正处于为英伟达(Nvidia)下一代AI加速器提供HBM3E验证的最后阶段。”

综合多家市场研究机构分析,预计今年SK海力士和三星的HBM市场占有率均为47%~49%,美光则有望达到5%。闪存市场分析师孙梦维认为:“盈利之后的原厂对市场占有率的要求开始提高,尤其是主要推动先进制程产能释放,预计第三季度在一定利润水平下,对市场占有率要求将进一步提高。”

HBM代际升级主要体现在数据速率和容量密度上,以满足下一代高性能计算应用的需求。

转为芯片部门下的一个常设办公室,以弥补其在2019年解散HBM业务和团队的市场误判。时下,三星与SK海力士已在HBM4研发上“短兵相接”,三星HBM团队计划于今年下半年量产HBM3E,并于2025年生产HBM4,2026年实现量产。

作为HBM市场后起之秀的美光也在积极布局HBM4,据了解,美光计划在2026年至2027年间,推出容量在36GB到48GB、12/16层垂直堆叠的HBM4产品,并于2028年推出带宽增加至2TB/s以上的HBM4E。

TechInsights市场分析总监Boris Metodiev表示:“2024年上半年的半导体需求喜忧参半,由于生成式人工智能需求激增,存储器和逻辑出现反弹。然而,由于消费市场的缓慢复苏,加上汽车和工业市场的需求回落,模拟、离散和光电子产品出现了轻度的调整。”

Metodiev还表示:“随着人工智能向边缘扩张,预计消费者需求将得到提振,下半年可能会出现全面复苏。此外,随着利率下降(为消费者提供更多购买力)和库存下降,汽车和工业市场预计将在今年下半年恢复增长。” (微文)

SK海力士开发出新一代移动端NAND闪存解决方案

本报讯 存储芯片厂商SK海力士日前在其官网宣布,他们已经开发出新一代移动端NAND闪存解决方案,将在今年第三季度开始量产并搭载于端侧AI智能手机。

从SK海力士在官网公布的消息来看,他们新开发出的用于端侧AI的移动端NAND闪存解决方案产品,名为ZUFS4.0,是当前业界最高性能的产品。

SK海力士表示,作为新一代移动端NAND闪存解决方案产品,ZUFS4.0专为端侧AI智能手机进行优化,按数据的特性来区分管理智能手机应用程序生成的数据,可以对不同用途和使用频率的

数据进行分区存储,提高操作系统的运行速度和存储设备的数据管理效率。

通过分区存储,智能手机应用程序的运行时间也得到改善。SK海力士提到,与现有UFS相比,ZUFS将在长期使用环境下,手机应用程序的运行时间提升了约45%,同时在存储读写性能下降方面实现了4倍以上改善,从而将产品的使用寿命提升了约40%。

在AI存储领域,SK海力士目前在高带宽存储器上已有很大的进展,随着ZUFS4.0的推出,他们预计在NAND闪存领域也将引领AI存储市场,巩固全球顶级AI存储供应商地位。(韩文)

铠侠公布最新季报 营收同比增长31%

本报讯 日前,日本存储芯片大厂铠侠公布了2023财年第四财季(2024自然年第一季度)财报,其营收同比大涨31%,净利润也实现了6个季度以来的首度扭亏为盈。

财报显示,受益于NAND Flash原厂减产、市场需求回暖及价格上涨,铠侠2024年第一季度营收同比增长31%至3221亿日元,为近七个季度以来的首度正增长;营业利润也由去年同期的亏损1714亿日元转为盈利439亿日元(前一个季度为亏损650亿日元),为近六个季度来首度扭亏为盈;合并净利润也由去年同期的亏损1309亿日元转为盈利103亿日元(前一个季度为净亏损649亿日元),为近六个季度以来首度盈利。

从铠侠的整个2023财年来看(2023年4月—2024年3月),其合并营收同比下滑16%至10766亿日元,合并营业亏损额达2527亿日元(2022财年为营业亏损990亿日元),合并净亏损额达2437亿日元(2022财年为净亏损1381亿日元),连续第二年陷入亏损,亏损额

本田和IBM联合研发下一代半导体和软件技术

本报讯 本田近日宣布,为实现未来SDV(软件定义汽车),解决处理能力、功耗、半导体设计复杂化等课题,公司与IBM签署了长期合作研发和共享下一代半导体、软件技术的备忘录。

本田在声明中表示,从2030年起,人工智能技术的使用预计将在整个社会中显著加速,使用这些技术的SDV有望成为移动出行的主流。与传统移动出行相比,SDV有望显著提高所需的处理能力和相关功耗,以及半导体设计的复杂性。

双方表示,为了解决这些问题并实现极具竞争力的SDV,必须具备自主研发下一代半导体和软

件技术的能力。随着汽车制造商在自动驾驶和先进驾驶辅助系统领域展开竞争,根据合作双方的预期,SDV的需求将会增加。

据了解,该谅解备忘录确定了双方联合研究的潜在领域,旨在大幅提高处理能力并降低功耗,本田表示将考虑联合研发小芯片等半导体技术。在软件技术方面,本田的目标是通过与硬件的协同优化来提高产品的性能并缩短开发周期,探索开放和灵活的软件解决方案。

本田汽车声明称,通过此次合作,两家公司旨在实现具有世界最高处理速度和节能性能的SDV。(艾文)

创铠侠前身“东芝内存”于2017年4月独立以来历史新高纪录。东芝(Toshiba)于2017年4月将以NAND Flash为主营业务的半导体事业分拆出去成立了“东芝内存”,之后“东芝内存”在2019年10月1日更名为“铠侠”。东芝目前持有铠侠约40%股权。

铠侠指出,2024自然年第一季度NAND Flash售价(以日元计算)较2023自然年第四季度上涨了15%~19%,为连续3个季度上涨。如果以美元计算的售价,则环比增长20%左右。NAND Flash出货量则环比增长5%~9%。

关于市场动向及后续展望,铠侠指出,随着客户库存水平正常化、带动PC及智能手机需求呈现复苏,今后期待On-Device AI登场,存储搭载容量增加以及PC操作系统更新带来的换机需求。

铠侠表示,AI普及以及存储搭载容量增加带动NAND Flash市场中长期呈现成长趋势的市场见解未发生太大变化。铠侠指出,将持续配合需求动向进行生产以及优化行销费用。(凯侠)

格芯任命洪启财为亚洲区总裁 将重点拓展中国的业务与战略合作

本报讯 5月20日,半导体制造商格芯(GlobalFoundries)宣布任命行业资深专家洪启财为公司亚洲区总裁兼中国区主席。洪启财表示,将以中国为重点,领导并发展格芯在亚洲市场的新业务与战略合作。

洪启财于2010年加入格芯,并担任过多项高级领导职务。就任亚洲区总裁及中国区主席前,他任职格芯首席制造官,管理格芯在全球的生产基地,并主管格芯在新加坡的所有业务运营。

同时,洪启财担任新加坡国立研究基金会(National Research Foundation in Singapore)董事会成员,并兼任国际半导体产业协会(SEMI)东南亚区域咨询委

员会主席。

格芯首席商务官Niels Anderskov表示:“洪启财拥有30多年行业经验,与亚洲半导体行业的很多领军人物建立了长期的良好关系,因此他是推动格芯在亚洲客户拓展和合作,特别是加速格芯在中国市场业务增长的理想人选。”

洪启财表示:“格芯差异化的技术解决方案与卓越的制造能力,特别是服务汽车等重要终端市场的能力,使我们在亚洲成为业务增长迅速的客户群体理想的合作伙伴。我们也在和很多跨国客户积极探索合作,通过高质量的制造和关键芯片技术,满足中国终端客户不断增长的需求。”(许子皓)