

# 英特尔硅光集成技术取得新进展

本报记者 张心怡

随着生成式人工智能快速升温，AI相关技术应用多点开花。而大体量的AIGC相关应用，几乎都植根于部署在数据中心的基座大模型上或部署在服务器集群、边缘服务器的模型上。由于大模型的参数量大，无法用单个服务器节点支撑，因此需要多个服务器节点乃至多个机架连接组成的集群。无论是应对大模型的高计算密度和高访存量，还是更大规模的处理器集群和服务器集群的部署要求，都需要更高的I/O带宽和更长的传输距离，以解决设备内部或设备之间的数据传输瓶颈问题。

在这种趋势下，半导体企业陆续开展硅光子及共同封装光学元件技术布局。顾名思义，硅光子技术集合了硅集成电路和半导体激光，用光学I/O取代电气I/O进行数据传输。相比以铜线互联的传统电子产品，硅光子技术支持在较远距离实现更快的数据传输速度。

在近日举办的媒体活动上，英特尔中国研究院院长宋继强介绍了英特尔在硅光集成技术的新进展。在6月举办的2024年光纤通信大会期间，英特尔团队展示了完全集成的OCI(光学计算互连)芯粒，在AI基础设施中实现了光学I/O封装。该芯粒目前单向支持64个32Gbps通道，传输距离达100米，有助于满足AI基础设施对更高带宽、更低功耗和更长传输距离的需求。

据介绍，OCI芯粒主要由PIC(硅光子集成电路)和EIC(电子集成电路)组成。PIC包含带有片上密集波分复用激光器，能够在光纤上复用多个波长的光，以及半导体光放大器，用于放大激光功率以满足传输所需的质



量和距离。EIC主要用于控制硅光子集成电路和连接主机。这颗芯粒还可以和CPU、GPU或SoC等计算部件封装在一起，提升了可扩展性和性能优化空间。“未来基于OCI，能够形成不同种类的计算+互连的芯片种类，适应更多应用场景。”宋继强说道。

宋继强向记者表示，硅光集成技术有两个优势：一是可以用半导体，特别是硅去发光和检测光，这意味着其生产流程可以与硅工艺集成起来；二是大规模的集成电路，也就是基于硅片，将硅与非硅的晶体管或者是其他电路形式做大规模集成。而英特尔的差异化优势在于，将高频率的激光发射器做在了晶圆上，再把硅的光放大器集成上去。这种片上激光器，不需要保持偏振

光特性不变的专门光纤，用普通光纤即可传输，在提升集成度的同时降低了成本。

据悉，英特尔已出货量超过800万个硅光子集成电路。目前，英特尔OCI的I/O接口芯粒实现了三个指标：一是在一根光纤里，可以分为8个波段进行稳定的传输；二是在每一个波段实现了32Gbps的数据传输速率；三是可以将8对光纤同时放在一起，互相之间不影响。

宋继强表示，未来将“迭代式稳步”提升OCI的I/O能力，第一步是在保持8种不同光波段的情况下，将数据的传输速率提升到64Gbps，使传输速度达到4Tbps。第二步是把光变成16个波段，使传输速度达到8Tbps。未来将持续向上演进，逐步提升带宽能力。

## NEO半导体推出3D X-AI芯片 神经网络性能提升100倍

**本报讯** 近日，在美国加利福尼亚州圣克拉拉举行的FMS 2024(内存和存储的未来)会议上，NEO半导体公司(NEO Semiconductor)创始人兼首席执行官Andy Hsu宣布，将推出一款改变游戏规则3D DRAM——3D X-AI芯片，其具有AI处理能力，将数据存储和数据处理结合在单个芯片中，将神经网络(ANN)性能提高100倍，功耗降低99%。

Andy Hsu表示，基于NEO的3D X-DRAM技术，3D X-AI模拟神经网络，包括用于权重数据存储的突触和用于数据处理的神经元，使其非常适合加速下一代AI

芯片和应用。3D X-AI可替代高带宽内存(HBM)，有望显著推动AI芯片设计和AI工作负载优化。

据介绍，单个3D X-AI芯片包括300层3D DRAM，容量为128Gb，以及1层神经网络电路，具有8000个神经元，每个芯片支持高达10TB/s的AI处理吞吐量。使用12个堆叠HBM封装的3D X-AI芯片，只需一个3D X-AI芯片，即可将其容量和性能提升12倍，达到1536Gb(192GB)容量和120TB/s的处理吞吐量。

“典型的人工智能芯片使用基于处理器的神经网络。这涉及结合高带宽内存来

模拟用于存储权重数据的突触和图形处理单元(GPU)来模拟用于执行数学计算的神经元。性能受到HBM和GPU之间数据传输的限制，来回数据传输会降低AI芯片性能并增加功耗。”Andy Hsu说道，“具有3D X-AI的AI芯片使用基于内存的神经网络。这些芯片具有神经网络功能，每个3D X-AI芯片中都有突触和神经元。在执行AI操作时，它们用于大幅减少GPU和HBM之间数据传输数据的繁重工作负载。我们的3D DRAM极大地提高了AI芯片的性能和可持续性。”

(海文)

## 三星电机向AMD供应 用于超大型数据中心的高性能基板

**本报讯** 三星电机日前表示，已与AMD签订了超大规模数据中心用半导体基板FC-BGA(Flip Chip Ball Grid Array)的供应合同，并已进入大规模生产。

据业内消息，该产品将在三星电机的釜山工厂和越南新工厂生产。越南工厂是三星电机自2021年起投入金额超过1万亿韩元建设的FC-BGA专用生产基地。

FC-BGA是一种将半导体芯片与主板通过翻转芯片凸点连接的高密度封装基板，主要用于高性能计算(HPC)用半导体。

特别是随着大数据、机器学习等人工智能(AI)市场的扩大，预计对面积达2.25万平方米、至少运行10万台以上服务器的超大规模数据中心所需的高性能FC-BGA的需求将持续增长。

三星电机和AMD通过合作实现了两个以上的半导体芯片在一个大的半导体基板上排列，构建了集成系统。这种高性能基板对于中央处理器(CPU)、图形处理器(GPU)应用至关重要，提供了更大的面积和更多的层数，使得先进数据中心所需的高密度互连成为可能。

与普通计算机基板相比，数据中心用基板的尺寸大10倍，层数多3倍，确保了芯片间的高效供电和可靠性。

三星电机方面表示，通过创新的制造工艺解决了翘曲问题，确保了芯片的高产量。

三星电机计划通过与AMD的合作，引领高性能半导体基板市场。市场研究机构Prismark预测，半导体基板市场将从今年的15.2万亿韩元增长到2028年的20万亿韩元，年均增长约7%。

三星电机战略营销室副社长金元泽表示：“AMD是全球高性能计算和AI半导体解决方案领域的领先企业，我们已成为其战略合作伙伴。未来我们将继续投资于先进的基板解决方案，以满足不断变化的数据集中应用需求，为像AMD这样的客户提供核心价值。”

AMD全球运营制造战略副社长Scott Eiler表示：“与像三星电机这样的伙伴合作，确保了拥有提供未来几代高性能计算和AI产品所需的先进基板技术和产能。”

(星文)

## SK海力士预测

### 下半年面向AI服务器的存储器需求持续增长

**本报讯** 近日，SK海力士发布截至2024年6月30日的2024财年第二季度财务报告。报告称公司2024财年第二季度结合并收入为118.76亿美元，同比增长125%，环比增长32.1%；营业利润为39.54亿美元，同比增长289%，环比增长89.4%；净利润为29.79亿美元，同比增长238%，环比增长115%。

SK海力士表示，该业绩表现主要受益于HBM、eSSD(企业级固态硬盘)等适用于AI的存储器需求表现强势。DRAM和NAND闪存产品的整体价格持续上升，带动了收入环比增加32%。

据了解，DRAM方面，该企业从今年3月份开始量产及供应的HBM3E和服务器DRAM等高附加值产品的销售比重有所扩大。特别是HBM的销售环比增长80%以上，同比增长250%以上，这带动了该公司的业绩改善。

NAND闪存的销售以eSSD和移动端产品为主要增长拉动力，特别是eSSD的销售持续保持快速增长势头，环比增长约50%。SK海力士在财报说明中强调：“从去年第四季度起，NAND闪存产品的ASP

(平均售价)持续上升，实现连续两个季度盈利。”

SK海力士预测，下半年除面向AI服务器的存储器需求持续增长外，支持端侧AI的PC端和移动端新产品将会上市，高性能存储器销量也将随之增长，同时通用存储器的需求也将呈现明显的上升趋势。

当前，SK海力士已向主要客户提供12层HBM3E样品，计划在第三季度开始量产。

今年下半年，该企业将推出适用于服务器的32Gb DDR5 DRAM和用于高性能计算的MCRDIMM(将多个DRAM结合在基板上的模组产品)产品。在NAND闪存方面，海力士将扩大大容量企业级固态硬盘销售，海力士预测，此类产品收入将达到去年的4倍水平。

为了应对面向AI的存储器需求扩大，不久前开工的韩国清州M15X厂正在以明年下半年开始量产为目标进行建设工作。龙仁半导体集群目前正处在用地工程阶段，其第一座工厂将按原计划在明年3月开工，计划在2027年5月竣工。

(姬婷婷)

# 坚持纾困与培优两手抓 推动中小企业平稳健康发展

