



高通公司总裁兼CEO克里斯蒂亚诺·安蒙用一张图总结一年来高通在Oryon CPU上的成果

高通统一PC、手机、汽车三端芯片架构

本报记者 张心怡

在近日于美国夏威夷茂宜岛举行的高通骁龙峰会上，高通在智能手机和汽车平台拿出了三款骁龙Elite（至尊版）新品。其中最引人关注的，莫过于高通自研CPU架构Oryon的“开枝散叶”。在去年首次应用于高通AI PC芯片骁龙X Elite之后，Oryon CPU走进了此次发布的一款手机芯片平台和两款汽车芯片平台。加上高通自研的Adreno GPU和Hexagon NPU，高通补齐了自研SoC的最后一块拼图，也统一了PC、手机、汽车三条产品线的芯片架构。

在生成式AI的浪潮袭来之后，头部处理器厂商都不约而同地缩短了产品发布周期。

第二代Oryon CPU亮相

本次骁龙峰会上，高通公司总裁兼CEO克里斯蒂亚诺·安蒙（Cristiano Amon）没有参与三款新品的发布，但他花了相当的篇幅发布贯穿三款新品的主线：第二代Oryon CPU。

Oryon CPU的研发，可以追溯到2021年高通斥资14亿美元收购CPU设计公司Nuvia。虽然Nuvia当时仅仅创立两年，创始团队却拥有在苹果、谷歌等企业担任芯片架构师的经验，在公司成立一年后就推出了自研CPU架构“Phoenix”，以更低功耗实现了显著高于竞品的

峰值性能。

在2021年3月完成收购后，安蒙就表示高通计划将下一代CPU集成到智能手机、PC、数字驾驶舱、ADAS系统等产品线。而这一愿景，在之后3年的骁龙峰会上，一步步走进现实。

在2022年骁龙峰会上，高通公布了下一代名为“Oryon”的CPU核心设计，但并未展现具体细节。在2023年骁龙峰会上，Oryon CPU正式亮相，采用12个3.8GHz高性能核心，单核性能跑分领先竞品的同时，达到竞品峰值性能的功耗降低

了30%~70%，率先搭载于同期发布的新款PC芯片骁龙X Elite。

本次骁龙峰会发布的第二代Oryon CPU，性能相比第一代提升了30%，同等性能下功耗降低了57%。在单核、多核及达到相同峰值性能所需的功耗等参数领先竞品的同时，Oryon CPU在设备不插电的情况下，性能不会衰减。这一特点使Oryon CPU在不插电模式下，处理速度比竞品CPU快134%，也使骁龙X Elite在不插电模式下相比竞品处理器快90%。同时，由于第二代Ory-

on CPU将首次搭载到手机平台，因此相比竞品，Oryon CPU能够将性能和能效的提升兑现现在更小的封装中。

在生成式AI的浪潮袭来之后，围绕算力的竞争更加白热化，头部处理器厂商都不约而同地缩短了产品发布周期，从过去的2年一迭代走向1年一迭代。高通对于Oryon CPU似乎也有此打算。安蒙表示：“对于下一代Oryon CPU，我已经看到了一些高通工程团队所作的工作并为此感到振奋，请关注明年骁龙峰会的发布。”

自研CPU架构的利好，就是能够根据不同产品线的需求，对CPU进行微架构层面的调优和定制。

面向手机、汽车需求进一步定制化

自研CPU架构的利好，就是能够根据不同产品线的需求，对CPU进行微架构层面的调优和定制。“我们从头开始打造IP，Oryon CPU正是我们自研SoC的最后一块拼图。这种方式让我们拥有对CPU的完全自主权，能够直接定义和定制所有处理通道的配置，不再依赖外部（第三方）。”高通技术公司产品管理总监Manju Varma表示。

由于第二代Oryon CPU从PC平台走进手机平台，因此高通CPU设计团队首先面向移动平台进行架构调整。高通工程副总裁、Oryon CPU设计负责人Gerard Williams（曾为Nuvia创始人）在回顾团队这一年的工作时表示，他们大幅提升了Oryon CPU的时钟速度，同时降低了功耗，这是推动Oryon CPU走进移动终端的核心驱动力。同时，为了适应移动设备更小型的体积和

电池，团队也在减少芯片面积上下了功夫。

具体来看，骁龙8至尊版搭载的Oryon CPU采用了面向移动端设计的全新微架构。在核心设计上采用了2颗超级内核+6颗性能核心，不再采用能效核。超级内核引入全新的数据时序预取器，基于模式检测预取所有缓存行（CPU缓存存储数据的基本单位），并预测特定使用情况下哪些缓存行会被使用，以实现更高的频率（4.32GHz）和性能，同时提升在紧凑产品外形设计下的能效表现。6个性能核都经过调优，在保证能效的前提下运行密集型的应用程序。

对于不再使用能效核，高通的解释是，用乱序性能内核取代能效核，能够提供更好的性能和功耗组合，并配合超级内核实现更高的处理速度。

内存同样是提升CPU处理速度的关键。高通为每个CPU从集配置了24MB专用缓存。通常，增加一级缓存或二级缓存都会导致时延的增加。因此，高通在提升一级缓存容量的同时，采用缓存一致性协议，将时延保持在1纳秒内。同时，为了解决二级缓存扩容造成的时延，高通设计了与CPU内核同步的二级缓存，以降低时延和功耗。考虑到生成式AI高度依赖内存，骁龙8至尊版配备了目前最高速的LPDDR5内存。这样的微架构和内存系统升级，能够为用户带来更快的应用启动速度、更流畅的多任务处理以及生成式AI功能。

而骁龙座舱和RIDE至尊版使用的Oryon CPU，围绕车用平台进行了三方面的提升。一是处理能力较前一代车规平台的CPU提升

了3倍。二是强化了安全功能。三是面向汽车应用设计的适应性架构，能够支持先进舱驾功能和自动驾驶功能，也能够同一块SoC上同时支持舱、驾功能，这无疑契合了当前舱驾一体的融合趋势。

“在汽车平台，Oryon CPU的定制化聚焦安全系数，以确保汽车在智能驾驶等场景下的安全性，让GPU、NPU、DSP等IP模块在安全的基础上运行。”高通技术公司汽车、行业解决方案和云事业群总经理Nakul Duggal向记者表示。

当然，相比芯片平台本身，手机、PC和汽车厂商能够结合最新芯片能力设计出哪些功能，或许是消费者更为关注的点，而这需要高通与OEM及终端、车企继续加深合作，并通过高通AI Hub等模型库，为广大应用开发者提供更加趁手的工具和平台。

IC China 2024 将于11月18日在北京举行

本报讯 11月1日，第二十一届中国国际半导体博览会（IC China 2024）新闻发布会在北京举行。中国半导体行业协会副理事长兼秘书长张立、中国半导体行业协会执行秘书长王俊杰、北京赛迪出版传媒有限公司总经理宋波出席会议，分别介绍IC China 2024的举办意义、筹备情况、特色亮点，并回答记者提问。发布会由中国半导体行业协会专职副理事长刘源超主持。

发布会披露，IC China 2024由中国半导体行业协会主办，北京赛迪出版传媒有限公司承办，将于11月18日—20日在北京国家会议中心举办。自2003年起，IC China已连续成功举办20届，成为我国半导体行业的年度重大标志性活动。本届博览会将贯彻“集合全行业资源·成就大产业对接”理念，契合“创芯使命·聚势未来”主题，聚焦半导体产业链、供应链及超大规模应用市场，全景展现半导体行业的发展趋势和技术创新成果，汇聚全球行业资源，促进全行业合作交流，打造全球IC行业的顶级权威盛会。

据介绍，二十多年来，IC China作为中国半导体行业成果展示、交流合作、生态融合的平台，见证了中国半导体行业创新进步和前沿技术突破变革，助力中国半导体行业的发展。IC China 2024的举办具有四方面重要意义。一是发挥半导体全产业链配套功能。IC China 2024依托全产业链头部企业的吸引力和产业链整合能力，为国内半导体全产业链配套提供了良好的平台。二是打造“广开销路”新品展示平台。IC China 2024组织终端企业和经销商作为专业观众，并联合

汽车电子、家用电器等领域的行业组织，助力参展企业拓展客源和营销渠道。三是打造聚集各方要素的协同平台。IC China 2024瞄准产业的难点堵点，集合“政产学研金”多方力量，共同研讨交流，通过集聚各方要素，为半导体产业、企业赋能。四是发挥链接全球的国际合作作用。IC China 2024邀请了多个国家和地区的半导体行业组织代表参会，将组织多场主题边会活动，为国内企业出海和国际企业开拓中国市场提供参考信息与合作渠道。

据此，IC China 2024创新办会模式，呈现四大特色。一是聚合了全产业链。IC China 2024参展企业覆盖了全产业链主要环节，为产业精准对接搭建桥梁。二是会议规格升级。IC China 2024围绕产业链、地方、化合物半导体、新兴应用、半导体第三方服务、产教融合、国际展商、未来产业等主题设置展区，注重实际成果转化，将催生一批高水平、代表性项目达成合作意向。三是国际交流深化。IC China 2024邀请了美国、日本、韩国、马来西亚、巴西等国家的半导体行业组织代表、企业家、专家学者，以及国内外产业链主导企业代表，通过主题边会促进国际合作。四是推动产教融合与协同创新。IC China 2024将举办半导体产业前沿与人才发展大会、集成电路产教融合大型研讨会、“百日招聘”半导体专场活动等特色活动，推动产教融合，在半导体行业积极贯彻落实“就业优先”国家战略。

此次发布会得到社会各界关注，近百家新闻媒体出席了发布会。（张阔）

北方华创前三季度 净利同比大涨54.72%

本报讯 10月29日，北方华创发布了2024年第三季度报告。今年前三季度公司实现营业收入203.53亿元，同比增长39.51%；归属于上市公司股东的净利润44.63亿元，同比大涨54.72%；扣非净利润42.66亿元，同比大涨61.58%。其中，第三季度实现营业收入80.18亿元，同比增长30.12%；归属于上市公司股东的净利润16.82亿元，同比大涨55.02%。

北方华创表示，今年前三季度营收同比增长39.51%，主要原因是公司电子工艺装备收入同比增长46.96%，使得整体营业收入增加；净利润同比大涨54.72%，主要原

因是公司电子工艺装备收入增长较快，成本费用率下降，使得归属于上市公司股东的净利润增加；扣非净利润大涨61.58%，主要是由于公司电子工艺装备收入增长较快，成本费用率下降，使得归属于上市公司股东的扣除非经常性损益的净利润增加。

此外，北方华创前三季度经营活动产生的现金流量净额约为4.55亿元，同比下滑59.97%，其主要原因是公司为了满足订单需求，采购商品支付的货款增加，使得经营活动产生的现金流量净额出现下降。（北华）

天光半导体为神舟十九 配套多款高可靠产品

本报讯 10月30日凌晨，神舟十九号载人飞船发射取得圆满成功。神舟十九号载人飞船的成功发射，不仅是中国航天事业发展的又一里程碑，更是我们国家技术实力的又一重要体现，同时也激励着国产半导体科研生产水平的不断提高。

在此次发射任务中，天光半导体公司为神舟十九号配套了逻辑电路、驱动器、接收器、电平转

换及射频中频放大器等多款高可靠产品，涵盖计算机自动控制、导航、遥测、通信等领域。

作为国产集成电路供应商，天光半导体公司将坚持“拼搏、竞争、求实、创新”的天光精神，致力于技术创新和产品生产能力的提升，牢记“科技引领、产业报国”的使命，不断攻克技术难关，为航天事业提供更加优质、高可靠的产品和服务。（天光）

Yole:2024年全球晶圆设备收入将达1330亿美元

本报讯 近日，市场研究机构Yole Group发布的最新报告显示，目前半导体行业正处于强劲的上升轨道上，预计到2024年全球晶圆设备（WFE）收入将达到1330亿美元，同比增长19%。其中，83%来自设备出货，17%来自服务和支持。但是不同细分市场的增长将有很大差异。

Yole Group分析师将2024年WFE市场的增长归功于市场对面向生成式AI的DRAM/HBM和处理器的投资增长，而NAND Flash的资本支出仍然疲软，传统逻辑和专业市场的资本支出则面临潜在风险。在这种不确定的环境中，WFE供应商正在

通过使其应用组合多样化来维持或提高收入水平，从而应对不均衡的资本支出。

从更长的时间周期来看，Yole Group预计，到2029年，全球总体WFE收入预计将达到1650亿美元。其中，WFE出货量预计将增长到1390亿美元，年复合增长率为4.7%。这主要是由存储和逻辑芯片的设备架构变化驱动的。与此同时，服务和支持部分将产生270亿美元的收入，年复合增长率为3.3%。

值得一提的是，从区域营收来源看，在2023年和2024年，中国成为最重要的WFE设备发货目的地，占

WFE市场总收入的三分之一。从头部供应商营收来看，总部位于美国的公司传统上在营收方面整体处于领先地位。目前，WFE市场领导者是ASML、应用材料（AMAT）、泛林集团（Lam Research）、Tokyo Electron Limited（TEL）和科磊（KLA）。

Yole Group的分析师表示：“应用材料公司在2023年排名第二，并通过应用组合差异化实现了销售增长。与此同时，服务和支持部分有所下降（与NAND需求下滑相关），但在2024年出现了增长。最后，我们预计小型供应商的长尾收入将收缩。”（季文）

Gartner:2024年全球芯片市场将达6298亿美元

本报讯 市场研究公司Gartner 10月29日公布的最新数据显示，在人工智能需求的推动下，全球芯片市场将在2024年达到6298亿美元，同比增长18.8%，高于其一年前预测16.8%的增长率。不过，Gartner将其对2025年的预测，从同比增长15.5%下调至13.8%，因此明年市场总量将达到7167亿美元。

Gartner高级首席分析师Rajeev Rajput在一份声明中表示：“人工智能相关半导体需求的持续激增和电子产品生产的复苏推动了这一增长，而汽车和工业部门的

需求仍然疲软。短期内，存储市场和图形处理单元（GPU）将推动全球半导体收入。”

2024年的持续供应不足将导致NAND Flash价格在2024年上涨60%，但预计2025年平均单价将下降3%。随着2025年供应减少和价格不再上涨，预计2025年NAND Flash收入将同比增长12%。

Gartner表示，由于供应不足改善、高带宽内存（HBM）产量增长和不断增长的需求，以及DDR5 DRAM的价格上涨，DRAM供需将反弹，预计2025年

DRAM收入总额将达到1156亿美元，高于2024年的901亿美元。其中，最热门的产品可能是HBM，预计2024年和2025年的收入将增长28.4%和70%以上，分别达到123亿美元和210亿美元。预计2025年全球存储市场将创纪录地增长20.5%，总额将达到1963亿美元。

在AI的推动下，GPU收入预计为510亿美元，到2025年将同比增长27%。“然而，市场现在正在转向投资回报率（ROI）阶段，推理收入需要增长到训练投资的数倍。”Gartner分析师George Brocklehurst说道。（汇编）