

发布7纳米芯片，寒武纪为何低调？



本报记者 李佳师

1月21日，寒武纪在官网披露7纳米AI训练芯片思元 290 智能芯片及加速卡、玄思 1000 智能加速器的相关信息，并表示该系列产品于去年已经实现规模化出货。该披露信息中有几个关键词：“7nm”“量产”“AI训练芯片”。这不禁令人们感叹，寒武纪太低调了，其他AI芯片大厂发布一颗7nm芯片，往往要开各种发布会来告知全世界，而寒武纪只是在官网披露一下，没有任何其他动作。

异常低调的发布

思元 290 芯片可实现峰值算力提升4倍、内存带宽提高12倍、芯片间通信带宽提高19倍。

《中国电子报》记者在寒武纪创始人陈天石的微信朋友圈，看到了一条只有9个字的信息“思元 290 正式亮相”，转发的是寒武纪公司公众号内容，非常低调。但这并不影响资本市场的敏感嗅觉，当天寒武纪的股价上涨了 17.73%，收报 170 元。随着寒武纪的训练产品线集中亮相，寒武纪从训练芯片到云、边、端等业务布局已经清晰呈现。

思元 290 芯片以及这些加速卡为何物？在 AI 芯片市场处于什么样的地位？对于寒武纪与业界有什么突破价值？

人工智能芯片有很多种分类方法。按功能分类，分为训练芯片和推理芯片。按应用场景分类，分为云端和终端芯片。按照技术架构分类，分为冯·诺依曼架构和非冯架构芯片，例如类脑芯片等。这次寒武纪推出的思元 290 智能芯片是其首颗训练芯片。为什么要做 AI 训练芯片？AI 训练芯片市场有多大？其红利有多“红”？实际上，看看英伟达公司股价这些年在市场上的“一骑绝尘”，便能够理解。

随着 AI 算法突飞猛进的发展，越来越多的模型训练需要巨量的算力支撑，才能快速有效实施，可以说，算力是人工智能应用取得突破的决定因素。目前，英伟达在全球云端训练芯片市场一家独大，其 GPU+CUDA 计算平台被认为是最成熟的 AI 训练方案。此外，还有第三方异构计算平台 OpenCL+AMD GPU 以及云计算服务商自研加速芯片这两种方案。可

见，全球各路芯片厂商基于不同方案，都推出了针对云端训练的人工智能芯片。

从介绍来看，寒武纪推出的首颗训练芯片思元 290，采用台积电 7nm 制程工艺，集成 460 亿个晶体管，支持 MLUV02 扩展架构，全面支持 AI 训练、推理或混合型人工智能计算加速任务。相比此前寒武纪推出的思元 270 芯片，思元 290 芯片可实现峰值算力提升 4 倍、内存带宽提高 12 倍、芯片间通信带宽提高 19 倍，可提供更优性能功耗比，以及多 MLU 系统的扩展能力。

产品具有标志性意义

这次发布思元 290 之后，寒武纪的产品布局实现了“训推一体、端云融合”。

思元 290 芯片有三大关键点。

其一，是寒武纪采用 7nm 工艺制程。目前能够提供 7nm 制程芯片的设计企业，除了海思、英伟达等，寒武纪是少数进入这个阵营的企业。赛迪顾问集成电路产业研究中心分析师李秧认为，其工艺水平能够实现 7nm 制程，无论是在中国还是在全球，都是比较领先的，在技术上是具有标志性意义的。

思元 290 问世，网友“无关风月”在相关新闻下评论称：“寒武纪真是低调，发 7nm 制程的芯片，居然显得如此的平静。目前英伟达的 RTX 30 系也才用到 7nm 制程工艺。”工艺越高意味着复杂度也越高，能集成的晶体管数也越多，这次思元 290 晶体管数超过了 460 亿个。

网友“无关风月”提及的英伟达 RTX 30 系列，是去年 9 月由英伟达创始人兼 CEO 黄仁勋发布的，

RTX 30 基于 8nm 工艺由三星代工，黄仁勋认为 RTX 架构是英伟达有史以来提升最大的产品迭代。

其二，是寒武纪的多芯互联技术 MLU-Link。目前，随着算法模型的复杂度不断增加，为了构建更强大的计算平台，需要多芯片之间进行互联，对算力和训练的速度提出了更高的要求。而寒武纪最新推出的多芯互联技术，首次用在了思元 290 芯片上。有了这个技术，多卡矩阵就可以很轻松地链接起来。业内人士认为，每颗思元 290 的多芯互联总带宽能达到 600GB/s，这个数据与世界上最快的带宽速度相当。同时称，寒武纪的 MLU-Link 多芯互联技术应用起来很方便，尤其是对于算力中心而言，既高效又简洁。据介绍，寒武纪首款智能加速器玄思 1000 包含 4 片思元 290 智能加速卡，最大 AI 算力超过 4100 万亿次每秒（4.1PetaOPS INT4），一台玄思 1000 计算单元就足以替代一个小

型传统超级计算中心。

其三，是寒武纪对应用生态的支撑进一步提速。寒武纪称 Neuware 提供完善的开发工具包和社区支持，帮助用户在思元 290 芯片上进行方便、灵活的定制开发及部署工作。

软件平台和工具链是芯片公司构建生态的必备武器，英伟达之所以能够聚集起强大的 AI 生态，不仅仅是因为其芯片性能强大，与其软件平台和工具链的完善息息相关。某种意义上说，英伟达的“硬”实力得益于其“软”实力。

寒武纪称，其 Neuware 软件栈为思元 290 芯片提供完善的软件及应用生态，支持业界主流的 TensorFlow 和 PyTorch 等深度学习框架，用户不需要改变使用习惯，即可在思元 290 芯片上实现图形图像、语音、NLP、搜索推荐等多种应用的训练和推理。

星星之火可以燎原

（上接第 1 版）中国的苦难史、抗争史、改革史、开放史，浓缩在这座城市的街道、里弄、江畔、船舶……百川入海，星火燎原。一代人的觉醒，唤醒了一个民族。一代代人的传承，又将红色故事赓续出新的脚本。

毫鬓白发和垂髫少儿在文物前轻声交流着；几位外国友人看到激动处不约而同地鼓起掌来；一群少先队员雀跃而来，红色火种生生不息……在上海星罗棋布的“初心教室”里，前来瞻仰的人们带着这样的问题：中国共产党为什么能？马克思主义为什么行？一份份史料、一件件文物，也是一把把钥匙，开启了去解读、去探寻、去追问历史逻辑和精神密码的大门。

百年前，这个国家的主义很多，但那群青年只相信马克思主

义。寒武纪从终端 IP 场景开始，连续推出了寒武纪 1A、寒武纪 1H、寒武纪 1M 系列处理器，其后布局云端智能芯片及加速卡系列产品思元 100 和思元 270，在 2019 年推出基于思元 220 芯片的边缘智能加速卡。由此建立起覆盖云边端、训练、推理的产品矩阵，并打造了平台级基础系统软件 Cambricon Neuware。这次发布思元 290 之后，寒武纪的产品布局就实现了“训推一体、端云融合”。短短几年，寒武纪就成为了全面掌握通用型智能芯片及其基础系统软件研发和产品化核心技术的企业。

财报预测仍将持续亏损

现在是芯片企业最好的时代。埋头苦干，给出更优的产品，或许是芯片企业迎接这个时代最好的姿态。

此前，有机构希望邀请陈天石去大型会议做演讲，陈天石婉拒并表示，目前希望专注把产品和技术做好。或许，对于芯片公司而言，产品是面对市场的核心之剑，做好产品就是最好的宣传。

寒武纪财报显示，2020 年上半年营收 8720.34 万元，同比下降 11.01%，亏损 2.02 亿元，其中研发费用占营收比重为 318.1%，同比增加 182.69%。去年前三季度营收 1.58 亿元，同比上涨 42.97%，亏损约 3.095 亿元，研发费用 4.33 亿元，同比上涨 32.78%。财报称，公司预测第四季度仍将持续亏损，主要系公司目前仍属于大规模研发投入阶段。

现在是芯片企业最好的时代。埋头苦干，给出更优的产品，或许是芯片企业迎接这个时代最好的姿态。

如此涨价为哪般？

有人认为，这一轮芯片的涨价潮与此前不同。以往通常是几个品类因为“炒货”和“囤货”，导致部分芯片产品短期内供应不足，出现涨价现象，但很快就能得以恢复。而这次的涨价潮似乎来得更为猛烈，不仅是几个品类缺货，而是 MCU、GPU、显示驱动、TWS 耳机 SoC、功率 IC 等全线产品的缺货，最终导致了全产业链的涨价现象。

TrendForce 集邦咨询分析师曾冠玮向记者表示，现阶段，芯片涨价的“重灾区”涉及多个领域，其中，数据中心、网络通信、手机、个人电脑以及汽车等领域，都面临着芯片涨价的压力。

芯片涨价潮再次来势汹汹，如此涨价为哪般？

就终端芯片而言，业内专家普遍认为，新冠肺炎疫情初期，行业内部对未来的市场走向判断比较保守，终端厂商对备货均持谨慎的态度。但受疫情防控和居家隔离政策等因素的影响，数字化进程明显加

（本文作者为《人民日报》记者杜尚泽 刘士安 辛本健 曹玲娟）

市场需求日益旺盛，芯片产能紧张问题却迟迟得不到缓解，芯片涨价潮一浪接一浪。近期，晶圆代工厂联电拟再度调高报价，涨幅最高达 15%，这是公司在半年内第二次提高报价。与此同时，台积电汽车芯片子公司先进积体电路（VIS）也正在考虑最多达 15% 的涨价幅度。与此同时，下游封测厂日月光投控、京元电子等企业，也因芯片大量产出对封测需求大增，产能同步吃紧，纷纷传达出了涨价意向。

需求持续火爆 芯片将迎来涨价潮？

本报记者 张依依

涨价成厂商无奈之举

在半导体领域普遍供应吃紧、相关产品缺货严重的情况下，半导体市场中的供需失衡现象格外明显，逐步引发了“多米诺骨牌效应”。从晶圆涨价到材料涨价，再到 PCB（印制电路板）涨价、封测涨价，最后到芯片涨价……产业链上游材料、晶圆、封测等环节引发的芯片涨价潮正逐渐在整个行业中蔓延。

从 2020 年的第四季度开始，半导体领域的涨价现象就持续不断，有关芯片代工厂商 8 英寸晶圆厂产能紧张的消息不绝于耳。2020 年年底，产能紧缺的情况又延伸至 12 英寸晶圆。由于晶圆产能紧张、市场需求强劲，诸多厂商都提升了芯片代工产品的报价，芯片涨价也成为了各大厂商的无奈之举。

2021 年伊始，芯片涨价潮不仅没有呈现出将要褪去的迹象，反而有愈演愈烈之势。据不完全统计，2021 年以来，截至目前，已经涨价或者即将涨价的半导体厂商共有 40 余家。

其中，三星 CMOS 图像传感器已涨价 40%；瑞萨电子已经调涨部分模拟和电源产品价格，同时现有订单和新订单都将以新价格处理；意法半导体已经上调了全线产品的价格；华微电子已经把产品价格上调了 10%；瑞能半导体将提高部分产品的价格；赛灵思部分产品的价格将上调 25%；联电拟再度调高报价，涨幅最高达 15%；台积电汽车芯片子公司先进积体电路（VIS）也在考虑最多达 15% 的涨价幅度；得一微电子嵌入式存储控制芯片的价格更是上调了 50%，成为众多产品中涨价幅度明显的产品。此外，Diodes、富满电子、Micro-chip 等多家厂商都已经发布涨价通知并采取了涨价行动，这波芯片涨价潮似乎势不可当。

“从 2021 年 1 月开始，多家芯片公司都已经提高了产品的出厂价格，普遍的上涨幅度在 10%~15% 这个区间。”赛腾微电子有限公司董事长黄继颇告诉《中国电子报》记者。部分研究机构预测，2021 年，所有芯片产品的平均涨价幅度将超过 20%，有些产品的价格甚至会直接翻倍。

如此涨价为哪般？

足够的产能则是芯片价格走势恢复平稳的另一前提。随着新冠肺炎疫情在全球范围内的逐步好转，国外厂商将逐步复工复产。基于此，黄继颇认为，在 2021 年第四季度，芯片的产能会有所释放，这有助于市场中的供需关系达到平衡，进而推动芯片产品价格的正常化。

步日欣也向记者表示，预计芯片的产能调配会逐渐达到均衡，甚至部分细分领域还会出现芯片供给过于饱和的状态。“2021 年中，芯片产品的价格会趋于稳定。”步日欣说道。

未来，在半导体产业的发展过程中，相关厂商除了要积极布局产能之外，对各个方面都应该保持理性。在接受记者采访时，步日欣指出，当整个产业处在一个充满“恐慌”的状态时，就会有囤货动机。任何产业都经不起不理性的抢货和备货，因此业界还需要时刻保持理性，避免因恐慌性囤货等因素引起的价格不合理上涨。

快，新冠肺炎疫情催生的“宅经济”让居家办公、远程教育和家庭娱乐渐成潮流，因此与这些领域相关的芯片产品价格都随之水涨船高，特别是手机、个人电脑等领域的芯片产品。

除了庞大的市场需求外，终端芯片产品的涨价现象与芯片产能和市场氛围也有一定关系。黄继颇指出，由于新冠肺炎疫情在国外肆虐，国外 FAB 的产能自然会受到影响。另外，业内出现了恐慌性备货的情况，大家都在拼命囤货，加剧了供需关系的紧张程度，最终导致了芯片产品的涨价。

就汽车芯片而言，由于智能手机等终端产品的市场需求过于旺盛，半导体行业没有多余的产能供给汽车产业。此外，汽车所需的车规级芯片对技术要求极高，交付不良率要控制在百万分之一，复杂的生产流程和严苛的标准也是导致汽车芯片短缺、涨价的一大原因。

目前，芯片涨价的细分领域还集中在收益性较低的小品类产品上面。创道投资咨询总经理步日欣进一步解释道，因为这些产品的市场规模不大，且利润较为微薄，所以晶圆代工厂在产能下降的情况下，会优先供应市场规模大且收益性高的产品，如消费电子类处理器芯片等。受“优先供应”的影响，小品类芯片的产能会受到挤压，当供给跟不上时，涨价就成了必然结果。

价格走势或于年内平稳

迄今为止，本轮芯片涨价潮仍在持续，可能会在行业中产生一些传导效应。曾冠玮指出，芯片涨价的传导效应已经有所显现。具体来说，芯片涨价主要提升了终端电子的生产成本。曾冠玮认为，芯片涨价潮已经开始，预计在今年第三季度前不会有价格回跌的情况。

但是，鉴于充满很多不确定因素和突发状况的 2020 年已经过去，不少业内人士还是对未来芯片产业的市场需求和价格走势持有乐观态度，认为本轮芯片涨价潮将于年内得以终止。

芯片产品价格的回落需要供应与需求达到平衡。黄继颇预测，在 2021 年第四季度，芯片的市场需求会有所回落。他表示，这是因为从 2020 年第二季度开始，业内厂商都在超额备货，市场需求不足以支撑如此大的备货份额和持续性的备货行为，所以市场需求会从高点渐渐回落。

足够的产能则是芯片价格走势恢复平稳的另一前提。随着新冠肺炎疫情在全球范围内的逐步好转，国外厂商将逐步复工复产。基于此，黄继颇认为，在 2021 年第四季度，芯片的产能会有所释放，这有助于市场中的供需关系达到平衡，进而推动芯片产品价格的正常化。

步日欣也向记者表示，预计芯片的产能调配会逐渐达到均衡，甚至部分细分领域还会出现芯片供给过于饱和的状态。“2021 年中，芯片产品的价格会趋于稳定。”步日欣说道。

未来，在半导体产业的发展过程中，相关厂商除了要积极布局产能之外，对各个方面都应该保持理性。在接受记者采访时，步日欣指出，当整个产业处在一个充满“恐慌”的状态时，就会有囤货动机。任何产业都经不起不理性的抢货和备货，因此业界还需要时刻保持理性，避免因恐慌性囤货等因素引起的价格不合理上涨。