

# 马斯克的“星际晶圆厂”，可行之举还是工程幻想？



本报记者 张心怡

在智能体轮番炸场之后，以每年超过1太瓦当量的算力产出为目标（涵盖逻辑芯片、存储与封装）的TERAFAB晶圆厂，作为工程实现能力的代表，成为今年科技圈的又一剂猛药。作为马斯克团队“走向星际文明”蓝图的重要一环，TERAFAB将为Tesla（特斯拉）和SpaceX生产芯片，目标产能是全球现有晶圆产能的50倍，其中80%用于太空，20%用于地面。

更值得注意的是，该晶圆厂将构建包含光刻掩模版制造、芯片制造、芯片封测、芯片设计的“递归循环”，且以每天（滞后不超过7天）的频率迭代新的芯片设计。这一打破“历史经验”的超规格芯片制造厂，具备必要性和可行性吗？

## 仅供自用，需要那么大产能吗？

结合马斯克在TERAFAB发布会（北京时间3月22日上午）的演讲及X平台回复网友的信息，TERAFAB包含两座晶圆厂，设计生产不同的芯片。一类是针对边缘推理进行优化的芯片，主要用于Optimus机器人和汽车，尤其是Optimus（马斯克预计人形机器人的产量会是汽车产量的10倍到100倍）；另一类是为太空设计的高功率芯片，会面向高能离子、光子以及电子积聚等来自太空环境的干扰进行优化，并尽量减少散热器重量。

从目前的规划来看，TERAFAB是特斯拉和SpaceX的内部芯片厂，生产定制化的芯片，无对外服务计划。

## “递归循环”的芯片生产链路，实施起来有多难？

除了“行星级”的产能目标，TERAFAB的另一大看点是构建“递归循环”（recursive loop）的芯片生产路径。

马斯克表示，在TERAFAB这一先进制造晶圆厂中，将配备制造任何类型芯片（逻辑芯片、存储芯片）所需的所有设备，以及制造光刻掩模版所需的所有设备。“在单一建筑内，我们可以制造光刻掩模版、制造芯片、测试芯片，再制造下一个掩模版，并形成用于改进芯片设计的、极快的“递归循环”。”

熟悉SpaceX的读者，应该会觉得这种递归循环的理念有些耳熟，因为SpaceX在火箭制造上，就执行了类似的“迭代设计”理念。

据航天业记者艾瑞克·伯格所著的《马斯克：火箭革命》，要构建火箭这样复杂的系统，基本上只有两种路径可以选择——线性设计和迭代设计。其中线性设计会先设定初

始目标、明确所需条件，通过大量测试确保各子系统达标，再进行组装。其特点是进入开发阶段之前需要花费数年的时间进行工程策划，一旦进入制造阶段就很难对设计进行修改。而迭代式设计根据目标启动后，会很快进入概念设计、台架测试和原型开发环节，其精髓在于快速进行样机制造和测试，发现漏洞、再做调整。SpaceX团队在不断重复的过程中，捕捉主要缺陷，继而完善设计、持续迭代，开发出更为完善的产品，从而加快了项目开发过程。

但这套理念，要兑现在芯片制造——尤其是尖端制程的芯片制造上，可谓关卡重重。虽然马斯克没有在TERAFAB发布会提到该晶圆厂的工艺，但SpaceX董事Steve Jurvetson等行业内人士均透露，TERAFAB将以2nm为目标。这意味着这套“递归循环”的芯片链路不仅非常难，还极其昂贵。

在数据“用”的环节，数据安全与共享的矛盾突出。工业企业普遍对数据安全高度敏感，担心核心工艺数据、产线数据外泄。在“不敢共享”的顾虑下，跨企业、跨行业的数据流通难以实现。他表示，行业需要继续探索如何在保障数据主权的前提下实现“可用不可见”。

最后，在标准建设层面，王金桥表示，行业数据集的建设仍处于“各自为战”的状态，缺乏统一的标准规范，要建立统一的数据格式、质量分级、评测体系，推动形成规模效应和复用价值。

值得注意的是，部分数据集在建设时还存在与工业场景严重脱节的问题，脱离企业生产的实际需求，即便数据质量达标，也难以适配具体工艺和真实场景，无法真正赋能AI模型、提升生产效率，并可能造成数据资源和建设成本的浪费。

TERAFAB，要么我们就得不到这些芯片。”

对于具体需要什么芯片、多少芯片，特斯拉进行了更详细的计算，并发布在X平台上。该推文显示，为了尽可能多地捕获太阳能，需要每年将1亿吨的太阳能捕获设备送入太空，这需要“超大规模的能力”，包括将数百万吨级载荷送入轨道的发射能力、太阳能AI卫星，以及数百万台特斯拉Optimus机器人参与建设。

要实现上述目标，仅Optimus机器人就需要100~200吉瓦当量的芯片，再加上用于太阳能AI卫星的太瓦级芯片。这一需求已超过当前全球所有芯片制造商的总产能，甚

至超过它们到2030年的预计产能（基于现有扩产规划）。TERAFAB的使命是填补当前芯片产能与未来需求之间的缺口，使“星辰大海的未来”加速到来。

除了太空能源建设，马斯克对人形机器人市场规模的信心远超过电动汽车市场。人形机器人将成为TERAFAB所生产的边缘推理芯片的主要市场，以及TERAFAB“用于地面的20%产能”的主要构成部分。“地球上每年的汽车产量大约是1亿辆，我预计人形机器人的年产量将达到10亿到100亿台之间，这个量很大，特斯拉的目标是生产其中非常大的一部分。”马斯克说道。

另外，马斯克未必会照搬现有的芯片工艺流程。他在演讲中表示，TERAFAB的递归循环链路“会推动计算领域的物理学极限”，且团队“会尝试各种疯狂的东西”，也许会走向不一样的技术路径。如他所言，特斯拉起步时，很多人认为电动车难成气候，如今特斯拉每年生产200万辆电动车；SpaceX起步时，很多人认为重复使用火箭是不可能的，即便造出来也不具备经济上的可行性，如今SpaceX已经完成超过500次火箭着陆。无论TERAFAB能否如期交付，在研发和建设过程中的新成果，以及碰撞出的“无意识创新”，都值得业界期待。

最后是巨额的流片成本。据业内测算，2nm流片一次的金额高达

针对上述问题，业内人士普遍认为，各方应各司其职，协同助力高质量数据集发展。鼎捷数智执行副总裁刘波表示，希望平台机构以可落地的智能工具为核心，打造全链路数据治理能力；龙头企业开放场景与数据资源，引领智能治理实践；中小企业主动拥抱轻量化智能工具，低成本实现数据质量提升。

## 多个场景有望跑出“标杆”

京东工业研发相关负责人向《中国电子报》记者表示，工业领域数字化转型具有“一米宽、百米深”的特点，“一米宽”即行业是由若干业务边界清晰的场景组成，“百米深”即每个场景背后都蕴含了大量的工业知识和机理，每个场景的数

从目前规划看，TERAFAB是特斯拉和SpaceX的内部芯片厂，无对外服务计划。

至超过它们到2030年的预计产能（基于现有扩产规划）。TERAFAB的使命是填补当前芯片产能与未来需求之间的缺口，使“星辰大海的未来”加速到来。

除了太空能源建设，马斯克对人形机器人市场规模的信心远超过电动汽车市场。人形机器人将成为TERAFAB所生产的边缘推理芯片的主要市场，以及TERAFAB“用于地面的20%产能”的主要构成部分。“地球上每年的汽车产量大约是1亿辆，我预计人形机器人的年产量将达到10亿到100亿台之间，这个量很大，特斯拉的目标是生产其中非常大的一部分。”马斯克说道。

TERAFAB将以2nm为目标。这意味着这套“递归循环”的芯片链路不仅非常难，还极其昂贵。

另外，马斯克未必会照搬现有的芯片工艺流程。他在演讲中表示，TERAFAB的递归循环链路“会推动计算领域的物理学极限”，且团队“会尝试各种疯狂的东西”，也许会走向不一样的技术路径。如他所言，特斯拉起步时，很多人认为电动车难成气候，如今特斯拉每年生产200万辆电动车；SpaceX起步时，很多人认为重复使用火箭是不可能的，即便造出来也不具备经济上的可行性，如今SpaceX已经完成超过500次火箭着陆。无论TERAFAB能否如期交付，在研发和建设过程中的新成果，以及碰撞出的“无意识创新”，都值得产业期待。

另外，马斯克未必会照搬现有的芯片工艺流程。他在演讲中表示，TERAFAB的递归循环链路“会推动计算领域的物理学极限”，且团队“会尝试各种疯狂的东西”，也许会走向不一样的技术路径。如他所言，特斯拉起步时，很多人认为电动车难成气候，如今特斯拉每年生产200万辆电动车；SpaceX起步时，很多人认为重复使用火箭是不可能的，即便造出来也不具备经济上的可行性，如今SpaceX已经完成超过500次火箭着陆。无论TERAFAB能否如期交付，在研发和建设过程中的新成果，以及碰撞出的“无意识创新”，都值得产业期待。

最后是巨额的流片成本。据业内测算，2nm流片一次的金额高达

龙芯中科董事长胡伟武：

## 自主算力融入云生态的速度远超预期

本报记者 许子皓

近日，龙芯中科联合浪潮数据在北京举办了浪潮数据服务器虚拟化产品联合发布暨战略签约仪式，会上推出了首款基于LoongArch架构原生开发的虚拟化产品InCloud Sphere龙芯版，深度适配龙芯新一代3C6000处理器。双方就此达成深度合作，将在技术研发、产品适配、市场推广等多领域携手发力，共同推动自主算力芯片与云虚拟化生态的深度融合，为党政办公、电子政务等国家重点领域筑牢高性能、高安全的云底座。

当前，数字经济已成为驱动国家经济高质量发展的核心引擎，算力作为数字经济的核心生产力，其技术自主化、供给可控化水平，成为衡量国家科技核心竞争力与数字基础设施安全的关键指标。在国家大力推动集成电路、基础软件等关键核心技术攻关的背景下，我国算力产业正从硬件单点突破，迈向“芯片-软件-生态”全链条协同创新的新阶段。云计算作为算力释放的核心载体，虚拟化技术作为云基础设施的底层核心，其自主研发与产业化应用成为算力底座建设的重要抓手，国内虚拟化技术历经多年研发积累，已实现从跟跑到并跑的跨越，为自主算力芯片与云虚拟化生态的深度融合奠定了坚实的产业基础。

龙芯中科董事长胡伟武表示，自主CPU的价值不仅在于芯片本身，更在于构建完整、可控、可持续的自主信息技术生态，虚拟化作为云基础设施的核心底层软件，是龙芯LoongArch生态走向成熟、支撑关键行业核心业务的关键一环。

据了解，作为此次合作的核心硬件支撑，龙芯3C6000系列处理器整体性能对标英特尔至强三代产品，推出16核、32核、64核三款型号，还搭载龙链技术以降低跨片核间通信延迟。凭借全自主研发优

势，3C6000系列实现极致性价比，相同性能下成本较同类产品低一倍，成为云计算、云存储、智算中心等场景的优质选择，也为此次虚拟化产品的研发落地奠定基础。

浪潮数据与龙芯中科此次完成了底层操作系统全栈适配，在系统内核、Hypervisor、硬件驱动、安全组件等核心环节实现深度调优，全面兼容计算、存储、网络等常规虚拟化能力。产品依托龙芯LoongArch架构的硬件级安全能力，构建起从芯片底层到上层应用的全链路安全防护体系，可满足党政办公、电子政务、公共服务等重点领域的高安全、高可靠需求。未来双方将持续整合资源，深化LoongArch架构与虚拟化技术的融合创新，进一步完善全栈云解决方案，推动技术成果在金融、能源、通信、交通等更多领域落地。

今年的政府工作报告提出“打造智能经济新形态”，要实施超大规模智算集群、算电协同等新基建工程，加强全国一体化算力监测调度，支持公共云发展。

在此方面，胡伟武表示，龙芯3C6000系列发布后，已在河南航空港智算中心、京津冀万卡级智算枢纽等项目中得到应用，替代了原有国外架构的算力底座，打造了全自主的高性能算力平台，此次又与浪潮联合推出InCloud Sphere龙芯版，面对国家算力基础设施建设的战略要求，龙芯将以性价比为核心切入点，从去年下半年起已经开始重点布局服务器与算力平台领域。

胡伟武将服务器分为通算、智算、存储三类，依托信创市场完成技术推广，并凭借生态壁垒低、性价比突出的优势，在存储服务器、AI算力等应用场景快速落地。目前国内虚拟化技术逐渐成熟，国内前十的云平台中，天翼云、移动云、京东云等头部平台均已实现龙芯适配，其他平台正在适配龙架构版本，自主算力在云生态的融入速度远超预期。

## Arm发布首款数据中心芯片

在扩展性方面，该芯片采用300瓦TDP设计，每线程独立核心，可在持续负载下提供确定性性能，避免降频与线程闲置。

在能效方面，支持高密度1U服务器机箱的风冷部署方案，单机架可支持多达8160个计算核心；同时也支持液冷系统，单机架可实现超过45000个核心的部署规模。

作为Arm AGI CPU的早期合作伙伴及联合开发者，Meta利用该代理式AI CPU优化其全系统应用的基础设施，并与其自研的Meta训练与推理加速器（MTIA）协同部署，从而在大规模AI系统中实现更高效的编排与调度。双方承诺将围绕Arm AGI CPU的多代芯片产品展开长期深度合作。

除Meta外，Arm已确认与Cerebras、Cloudflare、F5科技、OpenAI、Positron、Rebellions、SAP、SK电讯等企业达成进一步的商务合作。这些客户将在代理式CPU核心应用场景部署Arm AGI CPU，覆盖加速器管理、控制平面处理、云与企业级API、任务与应用托管等领域。

行业可以先行，他指出，这些行业具备三大特征，第一，产业链条长、场景复杂度高、生产数据密集，企业生产过程中产生的经验数据、记录数据、参数数据等海量数据，为数据集建设提供了丰富的原始素材；第二，数字化改造起步早、设备联网率高，已形成较为丰富的数据积累与应用实践，部分龙头企业已实现IT与OT数据的初步贯通；第三，对质量管控、设备运维、供应链协同、能耗优化等智能化的需求也十分迫切，龙头企业带动效应明显。

然而，试点推广过程也需警惕短期效应，兼顾长远发展。亚信科技智能制造专项负责人刘长水指出，试点项目应警惕“盆景化”风险，防止停留在“做示范”层面。刘波也表示，先行先试中还存在多方协同难、试点经验复用难、投入产出机制不清晰等现实困难，制约着数据集

规模化建设与价值释放。

从产业链全维度来看，工业高质量数据集的建设并非单一主体的获利，而能带动整个工业生态的协同发展。北京中科汇联科技股份有限公司董事长邵世学认为，高质量数据集建设有利于加速工业数据标准化、价值化进程，实现数据汇聚、治理、开放与共享。

上下游企业都有望借此迎来发展机遇。刘长水表示，对数据服务企业而言，随着数据集建设的推进，数据咨询、数据治理、数据标注等专业服务需求将大幅增长；工业软件与平台厂商将迎来需求爆发，相关行动的推进将催生一系列行业数据可信互联平台、数据资源库的建设需求；大模型与工业智能体厂商垂域大模型也将受此影响加速落地；制造业企业则可能通过数据赋能实现提质、降本、增效。

（上接第1版）

工业制造领域的形态多样，主要包括时序数据、图像、图纸文档、仿真数据等。中国科学院自动化研究所副总工程师、武汉人工智能研究院院长王金桥在接受《中国电子报》采访时指出，在数据“采”的环节，工业场景中的数据往往来自不同设备、不同协议、不同模式——视觉、声纹、振动、温度、时序信号等交织在一起，多模态、多源异构数据的采集与对齐难度较大，行业内还缺乏成熟的自动化采集工具和标准化接口。

在数据“集”的环节，王金桥认为，难点在于标注成本高、专业性强、长尾场景覆盖难。他特别强调，工业场景中真正有价值的往往是那些“小概率、大影响”的长尾事件，这类数据极度稀缺，靠真实采集难以满足模型训练的需求。