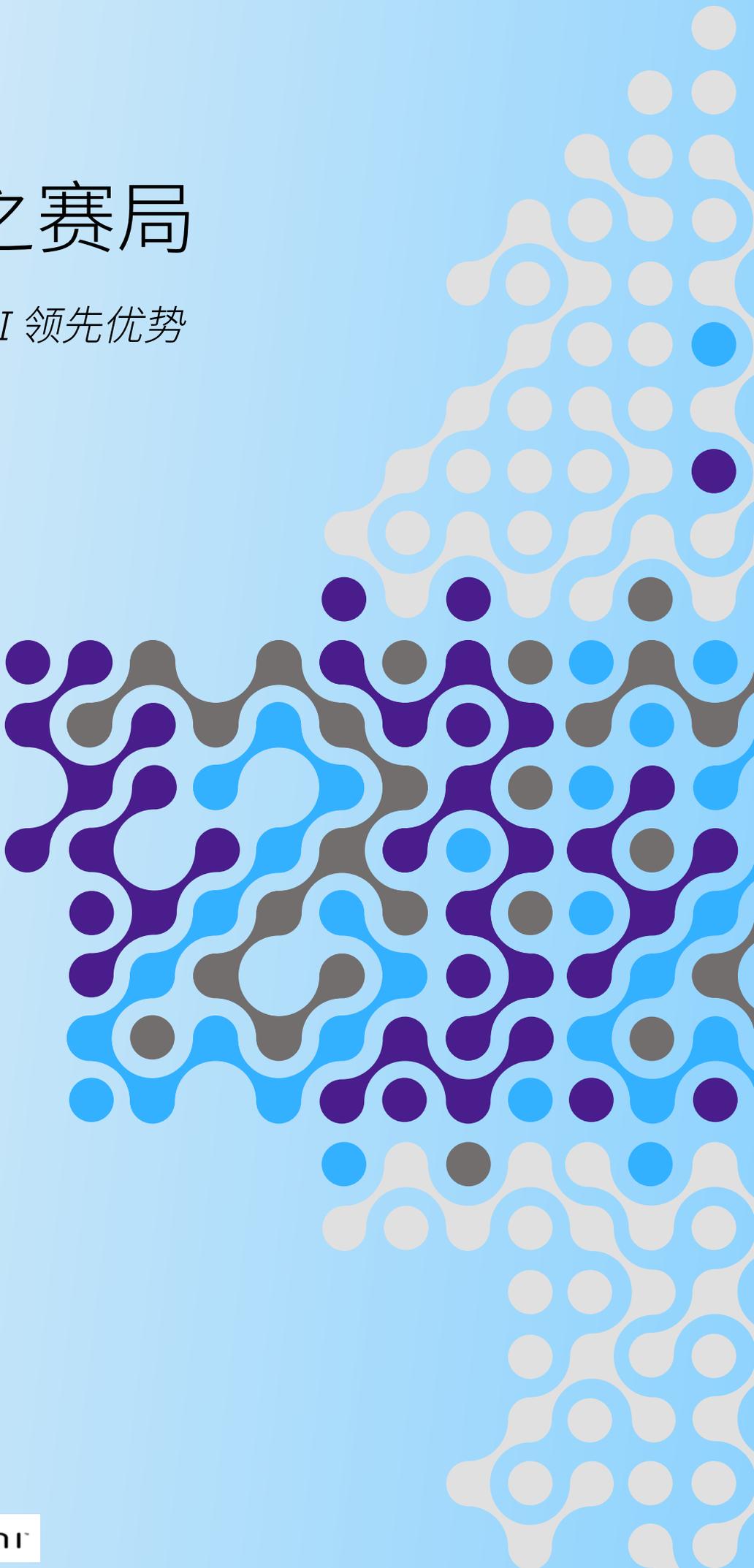


致胜硅之赛局

三大战略护航 AI 领先优势

中国洞察



IBM 如何提供帮助

在当今各行各业中，AI 战略已处于核心地位，而确保由高性能半导体所支撑的计算能力稳定供给，对于 AI 体系的韧性至关重要。IBM 提供成熟且经过市场验证的 CIM/MES 解决方案，帮助客户实现稳定高效的晶圆、芯片的制造和封装。在半导体运营方面，IBM 运用 AI 技术来提升产品开发效率并优化制造流程。通过结合先进的数据分析与预测性维护，IBM 助力客户提升产品良率并保障运营的可靠性。

如需了解更多信息，请访问：<https://www.ibm.com/consulting>

SEMI 如何提供帮助

随着 AI 需求激增、供应链承压，SEMI 正在为行业构建韧性平台。我们通过会员主导的系列倡议，聚焦四大关键优先领域：智能数据与 AI、供应链韧性、人才发展与可持续发展。SEMI 拥有超过 30 个技术社区——从先进封装到晶圆厂业主协会——促进专业知识共享，推动下一代技术创新。

SEMI 的全球性活动与专题会议，正在积极促成这份报告中所强调的供应商与买家的关键合作关系。我们提供市场情报、行业标准制定及概念验证 (PoC) 项目，加速创新落地。在 84% 的高管依赖政府激励政策的背景下，SEMI 亦致力于推动有利的政策环境与支持性框架的建立。

秉持“连接、协作、创新”的理念，SEMI 帮助成员多元化供应来源、构建区域生态系统，并保持 AI 发展的创新速度与活力。

序言

把握机遇，迎接产业升级

人工智能正以深远之力重塑产业格局，半导体作为其核心基石，亦迎来前所未有的机遇与挑战。电动汽车、机器人、传统电子设备迈向智能化，中国大陆半导体产业站上了转型升级的关键节点。

半导体芯片的市场需求呈爆发式增长。据国际半导体产业协会预测，2026年全球半导体设备销售额将攀升至1381亿美元，实现连续三年增长。¹中国已连续五年成为全球最大半导体设备市场，本土芯片产能接近全球三分之一。²规模背后，是对制造效率与稳定性的极高要求——尤其在大批量晶圆和芯片制造过程中，采用实时、集成、精准、可追溯的稳健和可靠的CIM/MES工厂全自动解决方案，是企业的核心竞争力所在。

半导体制造体系复杂而环环相扣，从原始晶圆制备、晶圆制造至封装测试，均需实现端到端的协同管理与精准运作。而在当前国际环境下，CIM/MES系统本地化已成为产业自主与供应链安全的关键举措。IBM凭借解决方案的开放性和端到端方案的整合实施能力，同本土伙伴合作，助力企业构建自主可控、高效稳健的“数字化工厂”能力。

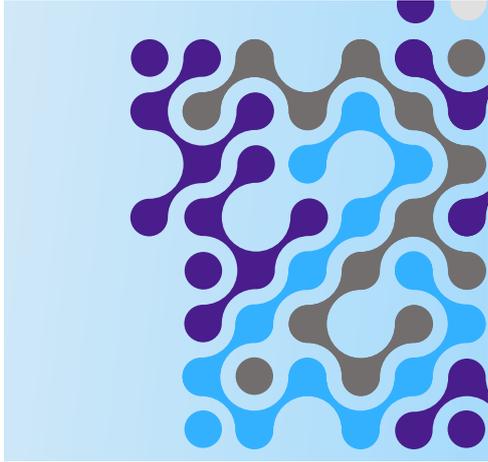
IBM商业价值研究院 (IBV) 研究显示，约九成的全球和中国供应商预计，至2028年，定制AI芯片、先进封装及新型计算架构将重塑全球半导体格局。本报告融汇了全球行业领袖观点与IBM实践洞察，从价值链重塑、生态构建与战略布局等维度，为企业提供可行指引。

前景可期，机遇已现。IBM愿与中国企业共同前行，推动半导体产业走向高质量与自主可控的新阶段。愿本洞察助您把握先机，迎接属于智能时代的产业新篇章。



IBM 咨询大中华区及韩国总裁 陈科典

2025年11月



摘要

- 芯片供应商难以跟上日益增长的 AI 需求**

AI 加速器芯片的需求预计到 2028 年将增长 50% 至 70%。83% 的芯片买家表示他们已经遇到供应瓶颈——而新的工业应用将进一步推高需求。
- 芯片买家正优先考虑本地化采购**

约 80% 的高管认为，获得本地生产的 AI 芯片、就近的 AI 人才资源和可访问的 AI 平台至关重要。然而，这一转变速度未能达到买卖双方的预期。
- 伴随 AI 的成熟，能效成为重中之重**

82% 的芯片买家正在寻求用于特定任务的专用芯片以优化能耗。近 90% 的供应商预计，对能够平衡性能、成本和能效的定制化系统级芯片 (SoC) 与芯粒 (Chiplets) 的需求将会增加。
- 下一代技术将有助于优化芯片性能**

88% 的芯片供应商预计，在未来三年内，包括光子计算、神经形态计算和量子计算在内的替代性计算技术将陆续涌现，以满足 AI 的发展需求。

目录

引言.....	4
第一章: 成为不可或缺的合作伙 伴.....	6
第二章: 打造区域生态系统增强 AI 韧性.....	12
第三章: 释放下一代 AI 芯片的最大潜能.....	18
附录: 芯片类型与应用指南	25

引言

以“缺”为机，赢在稀缺时代

算力跃升为全球硬通货，AI 加持下，将其需求推至高位。

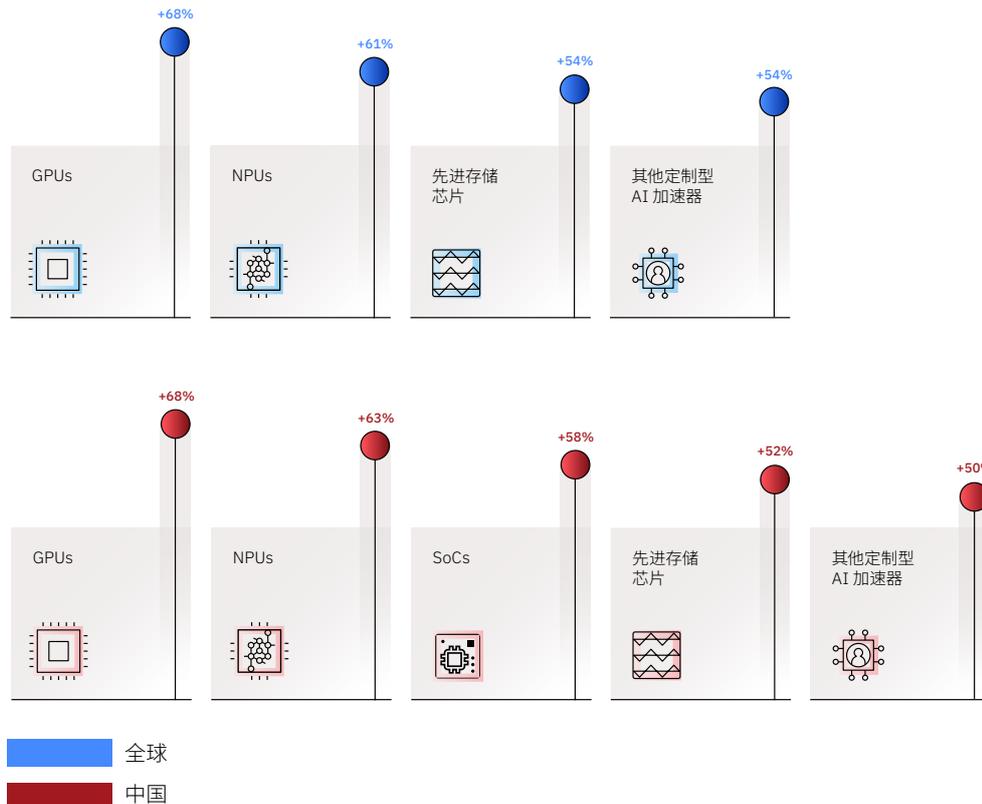
当企业竞相抢占人工智能驱动的商业新机遇时，其业务愿景依赖于纳米级半导体所提供的算力支撑。许多组织已将先进的 AI 芯片深度集成至其产品与服务中，然而，由于全球生产和资源瓶颈，芯片供不应求。

这一充满不确定性的半导体市场，将如何影响芯片买家与芯片供应商？在独家专有研究中，IBM 商业价值研究院 (IBM IBV) 与 Phronesis 携手合作，于 2025 年 5 月对 800 位芯片采购商与供应商高管（其中 100 位来自于中国大陆）开展了专项调研（详见第 27 页的“研究方法”）。调研显示，本已捉襟见肘的 AI 芯片市场将日趋紧俏。

受访芯片买家预计，到 2028 年，对最先进半导体——图形处理器 (GPU)、神经网络处理器 (NPU)、张量处理器 (TPU)、专用集成电路 (ASIC) 等 AI 加速器芯片——的需求增幅将高达 50% 至 70%（见图 1）。62% 的受访者表示，到 2028 年，高性能 AI 计算基础设施，将成为企业竞争优势的核心支柱。

图 1

先进 AI 芯片需求预计在未来三年内激增



问：预计未来三年内，贵组织半导体需求规模将呈现何种变化趋势（增加/减少百分比）？

百分比仅包含芯片买家的反馈。

有关完整的芯片类型术语表，请参见第 25 页的附录。

迄今为止，大部分需求源于虚拟世界的应用，例如大语言模型 (LLM) 训练和移动应用程序。随着 AI 技术成熟并更深入地融入物理世界——包括自动驾驶汽车、智能制造设备、机器人以及医疗器械等领域——其需求规模将发生根本性改变。针对特定领域的 AI 训练与边缘侧的嵌入式 AI 推理需求将显著增长。

目前，芯片供应链各环节的参与者——包括无晶圆厂 (Fabless) 和晶圆代工厂 (Foundry)、集成器件制造商 (IDM)、机械和材料供应商、电子设计自动化 (EDA) 以及知识产权 (IP) 供应商——均面临难以跟上需求的困境。约 83% 的芯片买家坦言，曾遭遇 AI 加速芯片断供或紧缺的问题。组织对智能体 AI 与各类 AI 场景的需求持续攀升，背后算力芯片的需求也随之走高，供需矛盾或将更加尖锐。

芯片买家陷入困境，75% 的买家表示依赖少数半导体供应商是主要的战略挑战。目前，掌握设计和制造最新的高性能 AI 芯片的公司屈指可数。新参与者正在进入市场，通常有政府支持——但技能和技术限制仍然阻碍着真正的竞争。

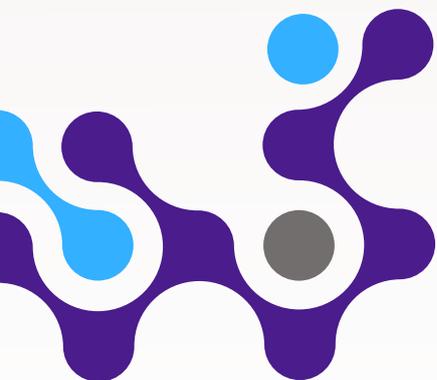
我们的研究揭示，解决方案需要创造性和新方法。在后续的部分中，我们都指出了能够帮助组织在高度波动的 AI 芯片市场中取胜的策略。在第一章节，我们将探讨如何立即重塑您与芯片供应商的关系，从而提升您在 AI 芯片交付名单上的优先级。第二章节深入探讨如何通过强化本地能力建设，提升 AI 系统韧性。第三章节，我们解释未来的先进 AI 如何推动开发更专用的芯片，从而在性能、成本和能耗之间取得更好的平衡。

这些策略的成熟节奏各不相同，但要构建真正的韧性，企业必须同时启动这三条路径，并从现在开始行动。



“AI 将像互联网和手机一样，通过训练和推理的加速计算，改变每个行业——从云到边缘，实现快速的模型开发，并影响全球每个市场。”

Nvidia
总监



第一章： 成为不可或缺的合作伙伴

卖方市场格局下，买方必须主动争取有利地位。首先需要洞察芯片供应商的痛点，再探寻协同破局之道。

受访供应商面临重重关卡。逾 80% 的供应商指出了四个关键难题：地缘政治紧张局势、投资限制、技术人才短缺及供应链技术瓶颈（见图 2）。有 89% 的中国大陆受访者认为，地缘政治竞争局势是当前最大的挑战，远超其他方面。对芯片买家而言，这些领域正是深化和改善与芯片供应链参与者关系的机会所在。

在地缘政治方面，芯片企业正在不同大洲建立新工厂，以应对不断变化的全球贸易环境和潜在的军事冲突——但目前，从启动生产到完成知识转移需要两年多的时间。在财务限制方面，81% 的全球芯片供应商和 78% 的中国芯片供应商表示新芯片技术所需的投资难以企及；74% 的供应商指出资本获取渠道有限。

芯片公司报告了关键技术的瓶颈，例如高带宽内存和晶圆基板。同时，61% 的受访供应商直言他们缺乏供应链管理技能，而仅半数表示他们拥有必要的 AI 专业知识。中国的受访芯片公司坦言，AI 算法与芯片集成是当前最大的技术瓶颈（见图 3A 和图 3B）。

“追求尖端科技，需持续且大量的投入。”

Toshihiro Awa
SUMCO
总裁

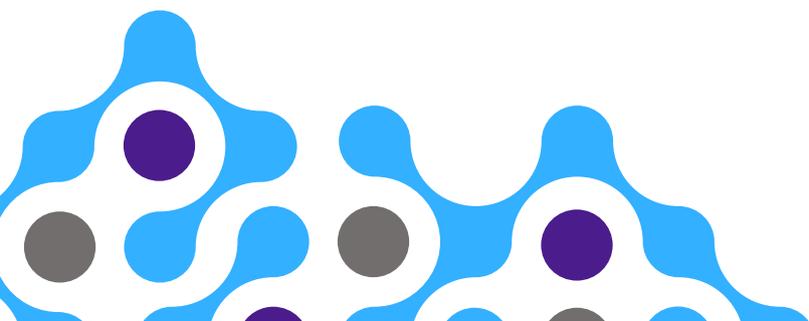
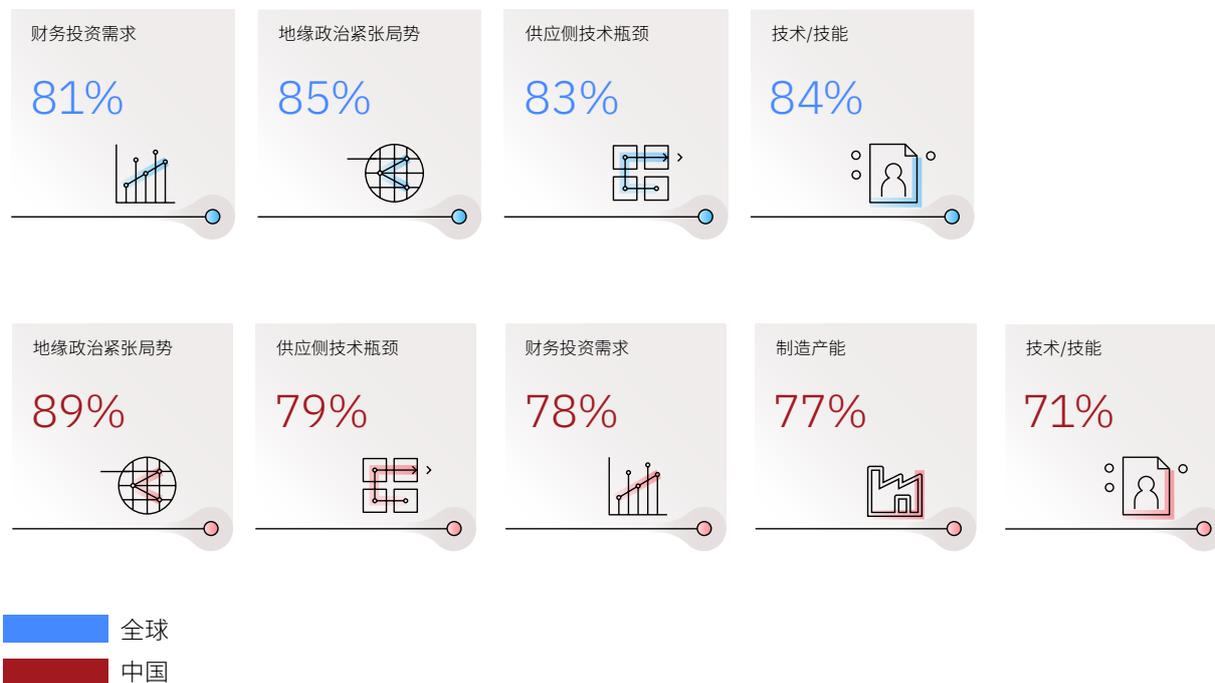


图2

芯片供应商多面承压、步步是关



问: 在满足市场需求与期望方面, 您认为以下因素对半导体行业构成的挑战程度如何?
(百分比为选择“存在一定挑战”和“存在重大挑战”的受访者比例合计)

芯片买家可以主动采取行动来应对这些挑战。目前, 80% 的芯片买家表示他们正在加强与芯片公司的战略联盟——84% 的买家希望建立新的联盟。

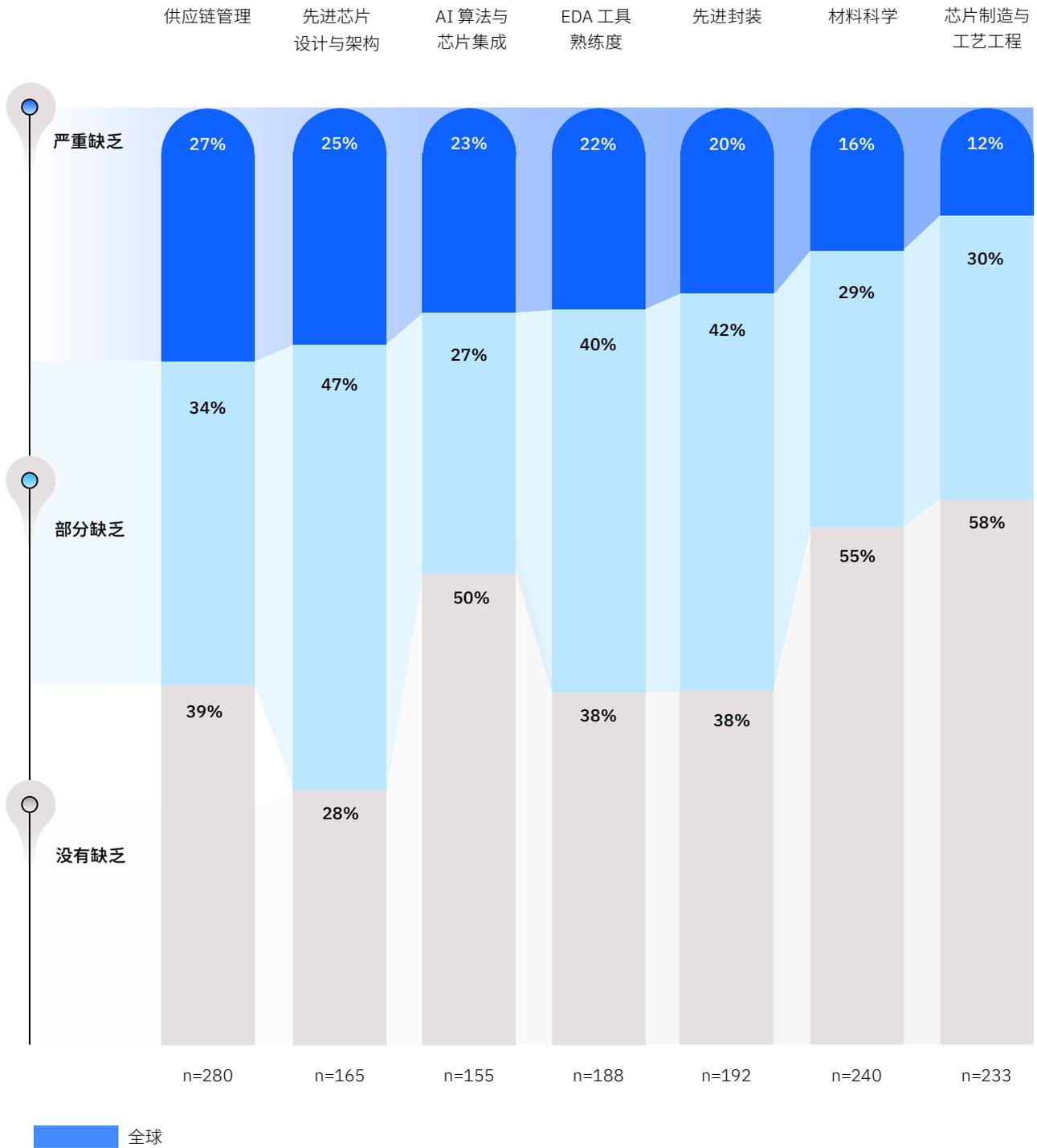
由于芯片买家掌握着驱动芯片设计的应用场景, 他们能够就优化半导体规格的技术要求提供宝贵见解。AI 越前沿, 芯片越专用, 买家可借助行业与场景专长, 在芯片开发中扮演更积极的角色。

“高端制造设备成本高, 却无平价替代方案, 这是全球业界共同困局。我们的目标清晰而坚定: 打造一个既高效务实又成本可控的解决方案。”

Hiroaki Takeishi
佳能工业集团
高级副总裁

图 3A

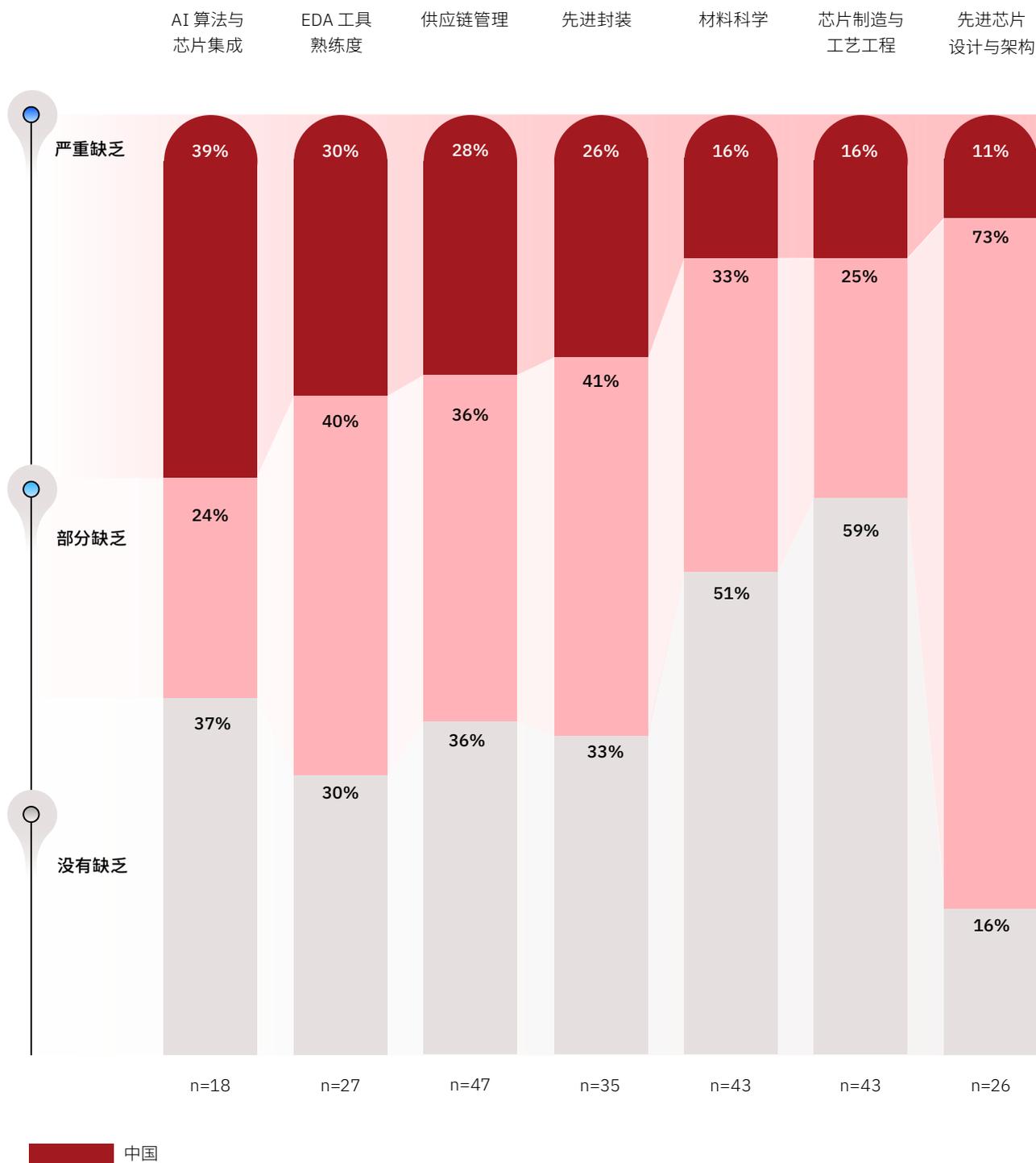
芯片供应商难以弥补关键技能缺口。



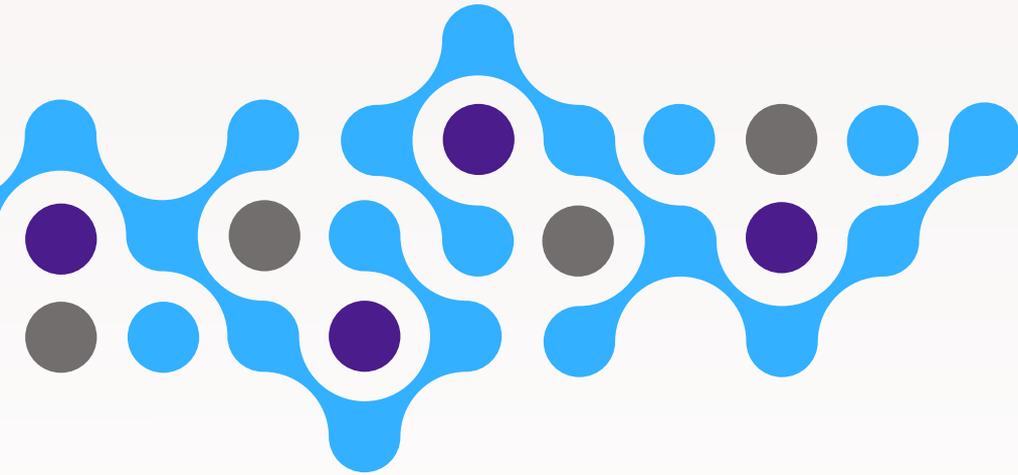
问：贵组织缺乏哪些关键的半导体技能？百分比表示认为其组织缺乏某项技能的高管比例。

图 3B

芯片供应商难以弥补关键技能缺口。



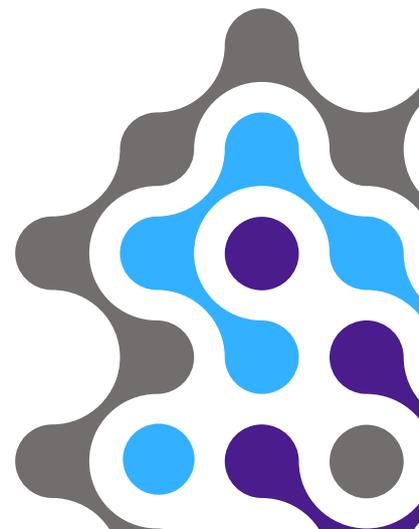
问：贵组织缺乏哪些关键的半导体技能？百分比表示认为其组织缺乏某项技能的高管比例。

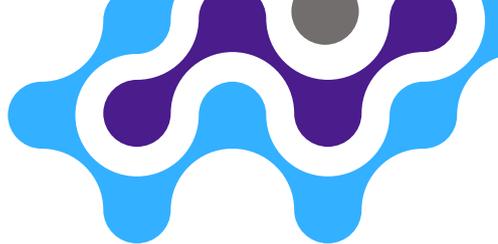


通过更深入的合作与长期采购承诺，芯片买家可以成为供应商的优选合作伙伴。例如，若在供应紧张时期，供应商往往更可能优先保障那些签署多年采购协议的客户，而非短期合作方。

这种合作伙伴关系还可以延伸至为新芯片运营提供本地支持、联合研发、新技术投资等领域。例如，买家可以利用自身在供应链管理方面的专业经验和网络资源，来赋能其芯片制造合作伙伴提升运营效率。他们甚至可以分享技术专长强化合作关系：65%的芯片买家表示他们正在积极扩展内部 AI 芯片设计能力。

通过创造性地提供资源，芯片买家可以成为其芯片供应商不可或缺的伙伴——从而使自己处于更有利的位置，既能获取芯片，又能使其长期 AI 战略更具韧性。





行动指南

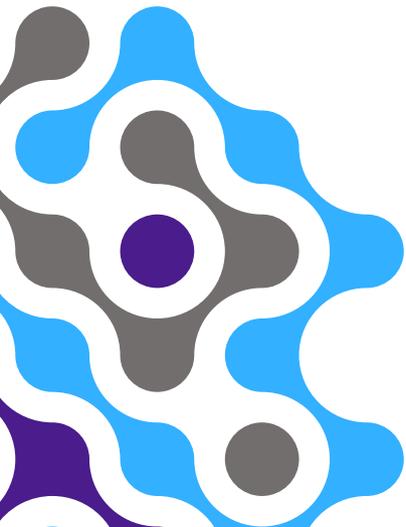
优先芯片合作伙伴关系，强化高效协作。投入资源，补齐半导体领域专业知识，以便与供应商更高效协同。识别行业瓶颈，探索多样化合作模式，补齐供应商在某些领域的知识短板。在半导体人才稀缺的环境下，探索多元且高价值人才的获取途径，例如通过AI代理填补人才缺口。

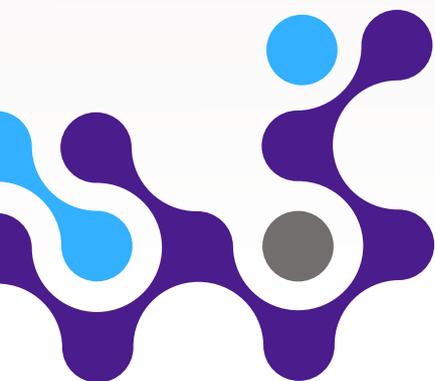
主导产品需求定义，深入芯片设计。定义清晰且有影响力的应用场景，帮助供应商优化芯片架构，实现成本、性能与能效的完美平衡。成为需求与合作节奏可预测的商业伙伴，帮助芯片合作伙伴更好地应对硅周期。

投资供应链透明化，筑牢风险防线。实现供应商多元化，并优化库存管理。借助AI场景演练，提前识别脆弱点。多预案并行，应对各种不确定性。确保供应链具备材料层级的可追溯性，以满足监管合规与社会责任要求。

“供应链透明在设备运营中至关重要。固态硬盘和半导体由众多组件与材料精密构成，任何单一组件的短缺都可能导致生产线停工。”

Koichiro Shibayama
KIOXIA Iwate
首席执行官





第二章： 打造区域生态系统增强 AI 韧性

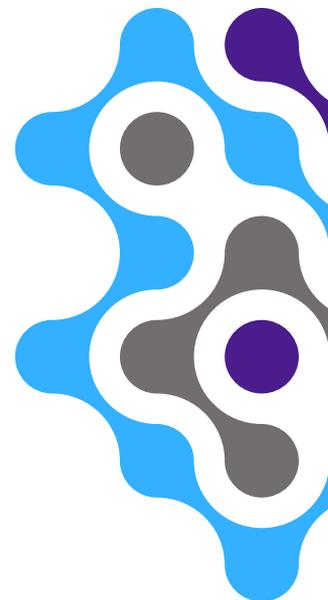
目前，只有少数地区能够供应最受市场需求的 AI 芯片及其关键支撑技术。这造成了供应链瓶颈，一旦出现扰动，难以迅速应对。在这种环境下，91% 的芯片采购商表示正在努力实现供应商多元化。

首要任务是实现地域多元化：约 80% 的受访高管认为，获得本地生产的 AI 芯片与加速器，对其所在地区至关重要；同样，约 80% 的高管认为，区域 AI 人才库很重要，而 77% 则强调 AI 平台的本地可及性。区域就近供给能促进整体生态系统的创新，同时减少易受干扰、关税敏感的全球链依赖。中国政府也在坚定不移地推动半导体国产化。通过政策扶持、资金投入与人才培育，助力突破核心技术壁垒，减少对进口的依赖。从芯片设计到制造封装，加速构建自主可控产业链，为行业筑牢安全根基。

然而，迄今为止，全球的区域化进展缓慢。至 2028 年，高管们预计区域化 AI 资源将满足重点领域的近半数需求（见图 4）。尽管公司认为高性能 AI 基建是关键竞争差异化因素，但 AI 芯片与数据中心硬件的本地化供给却未达预期。由此暴露出供应链上的要害薄弱环节。一旦缺少本地化资源，地缘政治动荡便会放大先进 AI 部署的风险。

我们的分析表明，有效的财务策略是加速 AI 计算硬件本地化采购的关键推动因素（见第 27 页的“研究方法”）。由于数据中心和 AI 芯片需要大量资本投入，早期的财务承诺将至关重要。

以半导体业务的财务需求为例，芯片制造厂建设的平均投资回收期为 75 个月，而新一代 AI 加速器芯片的建设平均投资回收期为 41 个月。同样，新建数据中心为 24 个月，新智能手机为 22 个月。

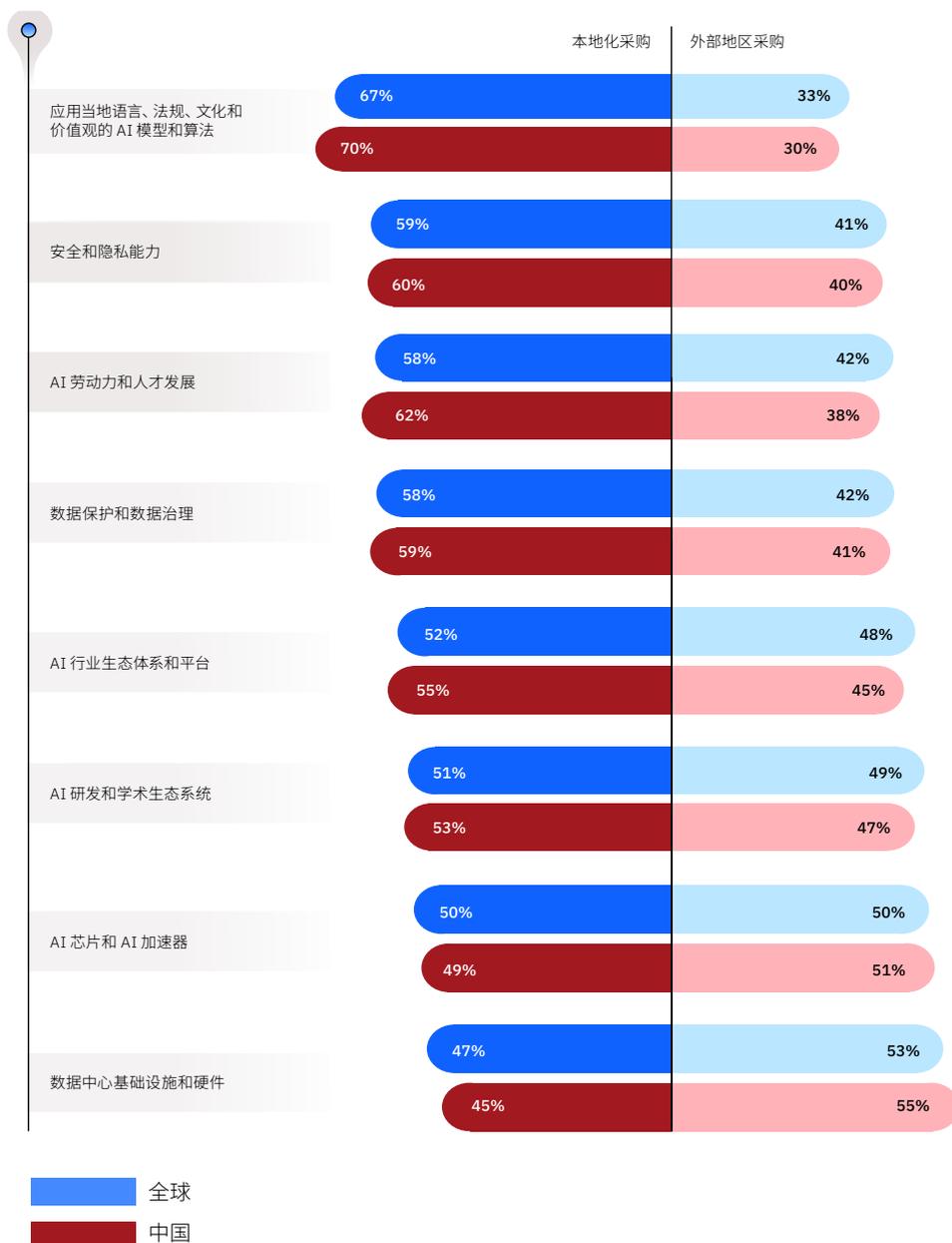


“地缘政治因素可能骤然冲击企业运营，
对公司构成重大威胁。”

Masato Goto
SCREEN Holdings
高级常务执行官

图 4

到 2028 年，区域 AI 资源对 AI 韧性至关重要。



问：您预计三年后贵地区有多大比例的 AI 需求将通过本地采购/能力来满足？

投资时间跨度的差异使得管理硅周期变得困难。芯片买家希望根据市场变化灵活调整订单，而芯片供应商则需要较长的前置时间和周密的规划，才能实现盈利性生产目标。这种供需之间的节奏错配，增加了合作的难度。因此，协调并统一财务预期，是确保合作伙伴关系成功的基石。

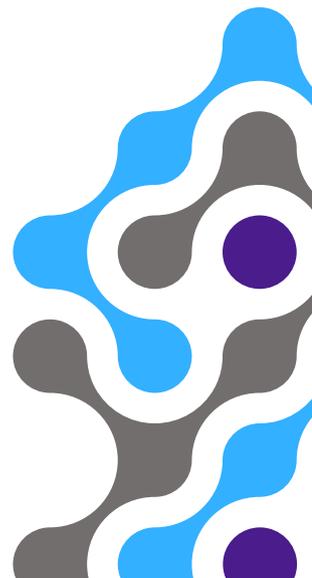
以特斯拉与三星的合作为例：双方近期达成合作，共同分担在美国建设新一代先进 AI 芯片制造工厂的财务负担。这项为期八年、总额 165 亿美元的合作协议约定，三星将在德克萨斯州建设工厂，用于量产特斯拉的 A16 专有芯片。³ 其他企业也有望受益于该工厂带来的美国本土产能。苹果已宣布将从该德州晶圆厂采购芯片，以推动其“美国制造”战略，同时规避潜在的关税风险。⁴

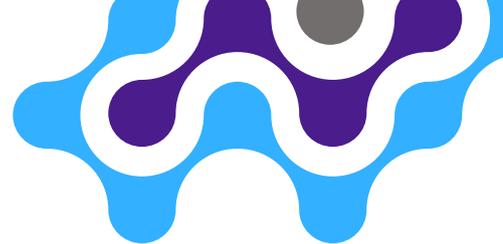
并非所有组织都具备特斯拉般的资源实力，这意味着组织可能需要组建合作联盟，吸引愿意出资合作的伙伴，包括对区域化芯片供应链感兴趣的行业协会。此外，争取政府机构的支持是另一个关键杠杆：84% 的受访高管表示，政府激励对组织成长与竞争力提升至关重要。

“昔日，高精工艺唯加州独揽；如今，工序已拓展至全球协作网络。”

Venkat Mattela
Ceremorphic
首席执行官

企业还必须思考如何利用 AI 驱动的创新释放新的收入来源，以支持区域半导体中心的持续发展。随着收入增长，企业就能投入更多资金，推动芯片供应链多元化。本土化能力的成熟需要时间，但一旦形成，将有可能打破现有格局，并增强整个区域生态系统的韧性。





行动指南



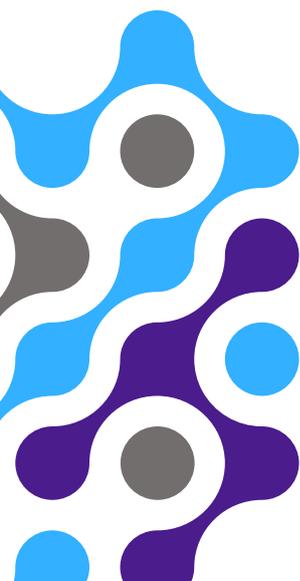
寻找具有共同 AI 发展愿景的区域合作伙伴。 发掘多元化的资金来源以支持区域芯片供应链建设，包括政府拨款、行业投资以及私营部门合作。充分利用 AI 赋能的收入流，为您的生态系统增值。

重塑采购、市场与合作战略，积极应对全球新变局。 提升或构建自身面向全球的情报能力，通过战略性投资再分配，为潜在的业务中断做好准备。让 AI 参与沙盘推演，使得规划更加准确和全面。

打造具备竞争力的 AI 算力，满足本地 AI 需求。 开发具有说服力且可靠的 AI 商业案例，依托于垂直整合的 AI 能力与您的专有数据。开拓新的 AI 收入来源，培育本地 AI 生态系统，携手跨越资本密集型芯片行业的最低投资回报基准线。

“半导体生产若向高成本地区转移，则必须降低对劳动密集型工艺的依赖。机器人技术与自动化将成为不可或缺的核心力量。”

Tsunetoshi Oba
OMRON 工业自动化公司
执行董事



观点

良性循环： AI 如何重塑半导体运营

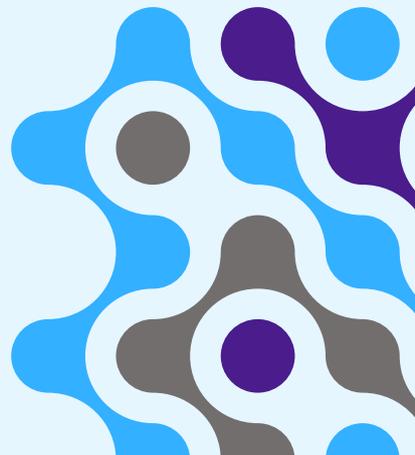
数十年的技术沉淀，筑就行业壁垒：顶尖组织的专业知识，正是前沿 AI 芯片研发复杂性与高成本之源。半导体行业持续优化运营的“独门打法”不断迭代，但离开特定公司与生态土壤就很难复刻。这些方法论往往由无数细微改进逐步积累而成，最终形成可观的竞争优势，成为企业的“护城河”。

长期以来，专业经验多未成文，以口传心授在组织内部流传，这种传承方式与高度数字化的产业特征格格不入。如今，半导体企业正借助 AI 优化内部流程，并在设计、制造、良率与质量控制、供应链和设备维护等环节提升知识传递能力——这已成为计划在新地区建立运营基地的企业的首要任务。AI 能够收集并分析历史记录，帮助支持和培训员工，从而确保宝贵的企业知识得以传承至下一代。

芯片供应商预计，未来三年 AI 将使运营成本降低 17%，并带动收入增长 32%。重点关注领域包括 AI 驱动的高级封装与组装优化，其中 73% 的供应商认为，这一领域的生产力有望提升 21%。

“在紧迫的时间约束下，AI 驱动的仿真已成为材料开发的关键手段。”

Toru Kimura
JSR 电子材料事业部
高级主管



半导体研发，尤其是先进工艺芯片的设计，也是 AI 能发挥价值的关键领域。由于 72% 的供应商表示在寻找先进芯片设计人才方面存在困难，因此 72% 的供应商转向 AI 优化电子设计自动化 (EDA)，预计可提升生产力 26%。

“我们立下雄心壮志：借力诸如数字化转型等先进解决方案，赋能数字化转型，将工厂生产效率提升至双倍。”

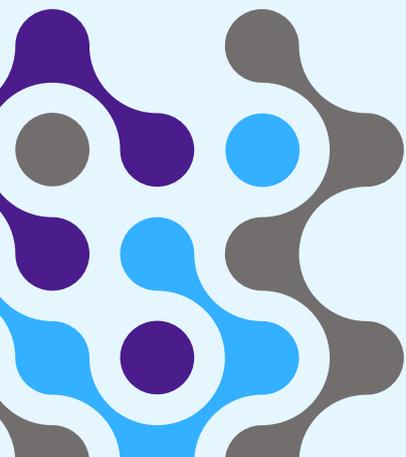
Kazunori Tsukada
KOKUSAI ELECTRIC
首席执行官

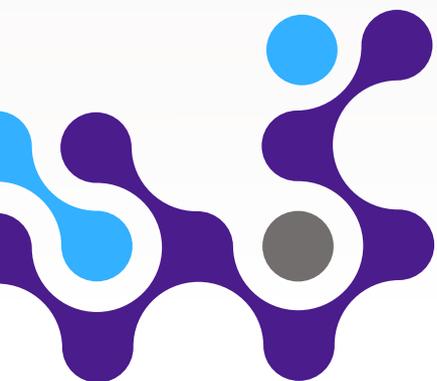
例如，领先的晶圆代工厂台积电就将“Fab Runs On Code”（晶圆厂以代码驱动）视为其竞争优势的秘诀。这一方法结合 DevOps、云计算和 AI，以提升良率和性能。⁵

各家公司正在全方位探索利用 AI 优化半导体运营的途径，从良率分析、视觉检测、预测性维护，到供应链与物流优化。能够成功利用 AI 精简运营的企业，会在即将到来的 AI 芯片需求激增中占据更有利的位置——而这波需求增长才刚刚开始。

“通过 AI 提升良率是当前的头号任务。哪怕只是提升 0.1%，也足以对业务产生显著影响。”

Toshihiro Awa
SUMCO 株式会社
社长





第三章： 释放下一代 AI 芯片的最大潜能

企业倾向于掌控自身的命运，尤其是在关键资源方面。然而，对于需求旺盛的计算芯片，这往往难以实现。但未来几年，这一局面有望改善——正如我们的研究显示，变革的种子已经在萌芽。

关键数据：83% 的芯片买家表示，技术进步已纳入其产品战略。这些企业需要持续高速的创新，以保持竞争力——因此他们正在考虑所有可能的选项。

专业化是芯片供应商跟上需求的一种方式。82% 的芯片买家正寻求针对特定任务的专用芯片，以优化能耗。通过专注于特定任务而非通用计算，他们能够在应用场景日益复杂的情况下，更好地平衡性能、成本与能效。

近 90% 的芯片供应商表示，其客户对定制 AI 芯片的需求将持续增长——主要是定制化的片上系统 (SoC)，即在单颗芯片上提供完整计算系统，而非依赖独立芯片或电路板。92% 的全球供应商和 93% 的中国供应商还认为，作为芯粒 (Chiplets) 发展关键推动力的先进封装技术将至关重要 (见图 5)。与现状相比，这一趋势正显著加速：目前仅有 24% 的 AI 加速器和 21% 的 SoC 是定制化的。

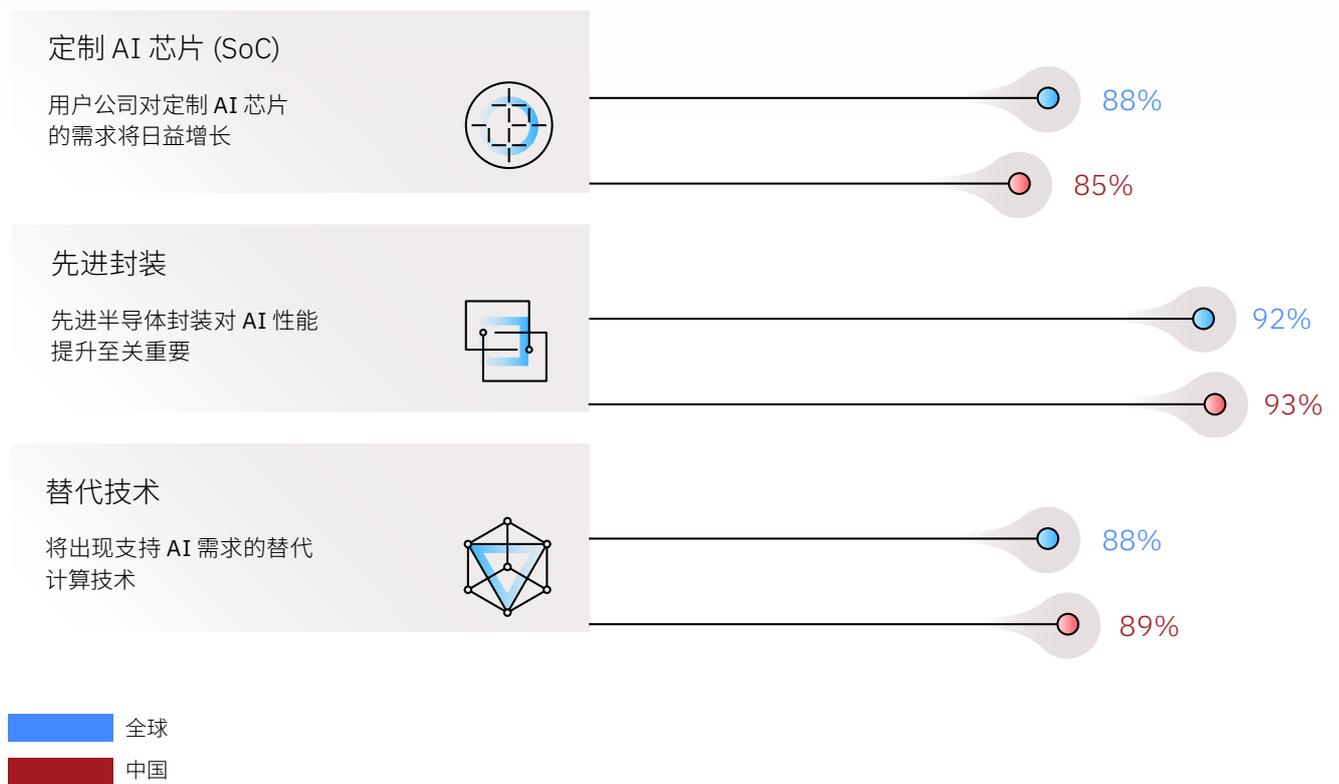
“我深信，面向特定场景的多品类 AI 芯片将迎来迅猛发展，其需求有望在 2027 或 2028 年前后迎来爆发式增长，进入行业发展的黄金期。”

Atsuyoshi Koike
Rapidus
首席执行官



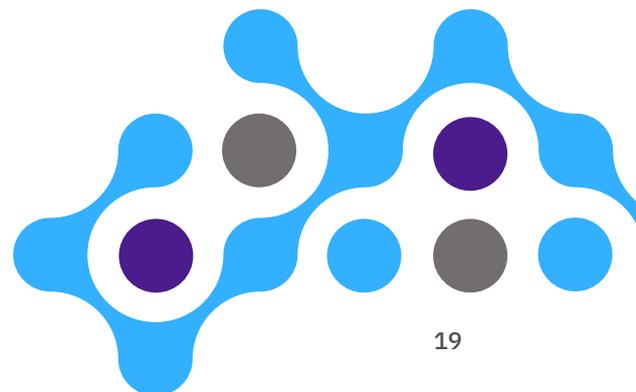
图 5

受访高管认为 2028 年前，三大趋势有望重塑全球芯片版图。



问：在三年内（至 2028 年），您对于以下陈述在多大程度上认同？

百分比显示选择“认同”或“高度认同”的高管比例。



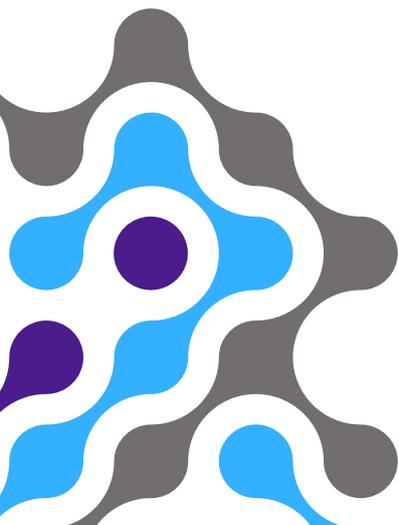
“基于一个基础芯片来开发定制 SoC 可能只需不到两年，而从零开始则需要更长时间。使用芯粒和标准接口可以显著缩短开发周期。”

Takao Fukuda
Astemo 战略部门产品规划
兼高级总监

短期内 SoC 将仍为主流选择，但随着封装技术的进步，组织正日益关注芯粒方案。芯粒具备多重优势：性能提升、能耗降低，架构更加灵活。SoC 采用单片设计，须以最尖端制程支撑全系统，然而，部分功能或无需如此极致工艺，难免大材小用。芯粒开启定制化新纪元，赋能精准优化，性能跃升。

在这种环境下，供应商必须为每一类应用、每一位客户，在性能、成本和能耗之间找到最佳平衡（见第 24 页“解决 AI 能源问题”）。供应商多管齐下，革新芯片设计、升级制造工艺，更以突破性技术引领变革。

未来三年内，88% 的全球芯片供应商和 89% 的中国芯片供应商预见，将会出现支持 AI 需求的替代计算技术，这包括支持芯粒发展的先进封装技术，以及诸如光学与光子计算、3D 堆叠架构等。





此外，还有一些突破性能力（见图6）。量子AI利用量子力学原理增强机器学习算法，可能使企业能够解决传统计算技术难以触及的复杂问题。人工通用智能（AGI）也被寄予厚望——即能够广泛理解、学习并跨多任务应用知识的AI。如果实现，AGI将为机器人技术、自然语言处理和计算机视觉等应用开辟新的可能性。

“对高性能 SoC 的需求虽然存在，但我们必须在性能与可负担性之间取得平衡。当成本超出客户可接受范围时，我们往往需要在规格上做出妥协。”

Honggul Jun
LG 电子汽车中控事业部
副总裁

然而，此类突破归根结底有赖于半导体技术的自身发展。83%的半导体供应企业高管表示，AGI能否实现，高度乃至极度依赖于半导体技术的进步。

尽管这些技术尚处萌芽阶段，但其商用可能比看上去更近在咫尺。那些能够与市场同频进化的组织，将在技术成熟之日抢占发展先机，赢得巨大优势。

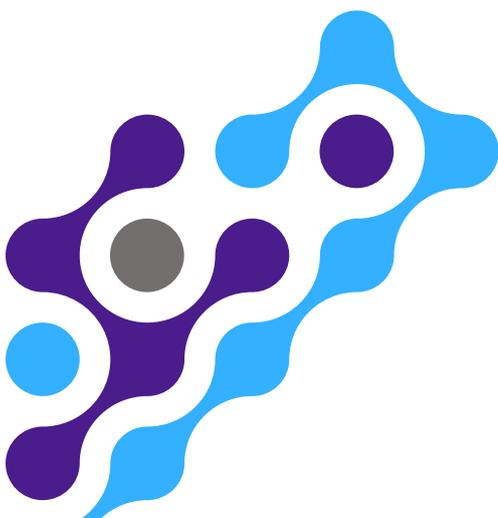
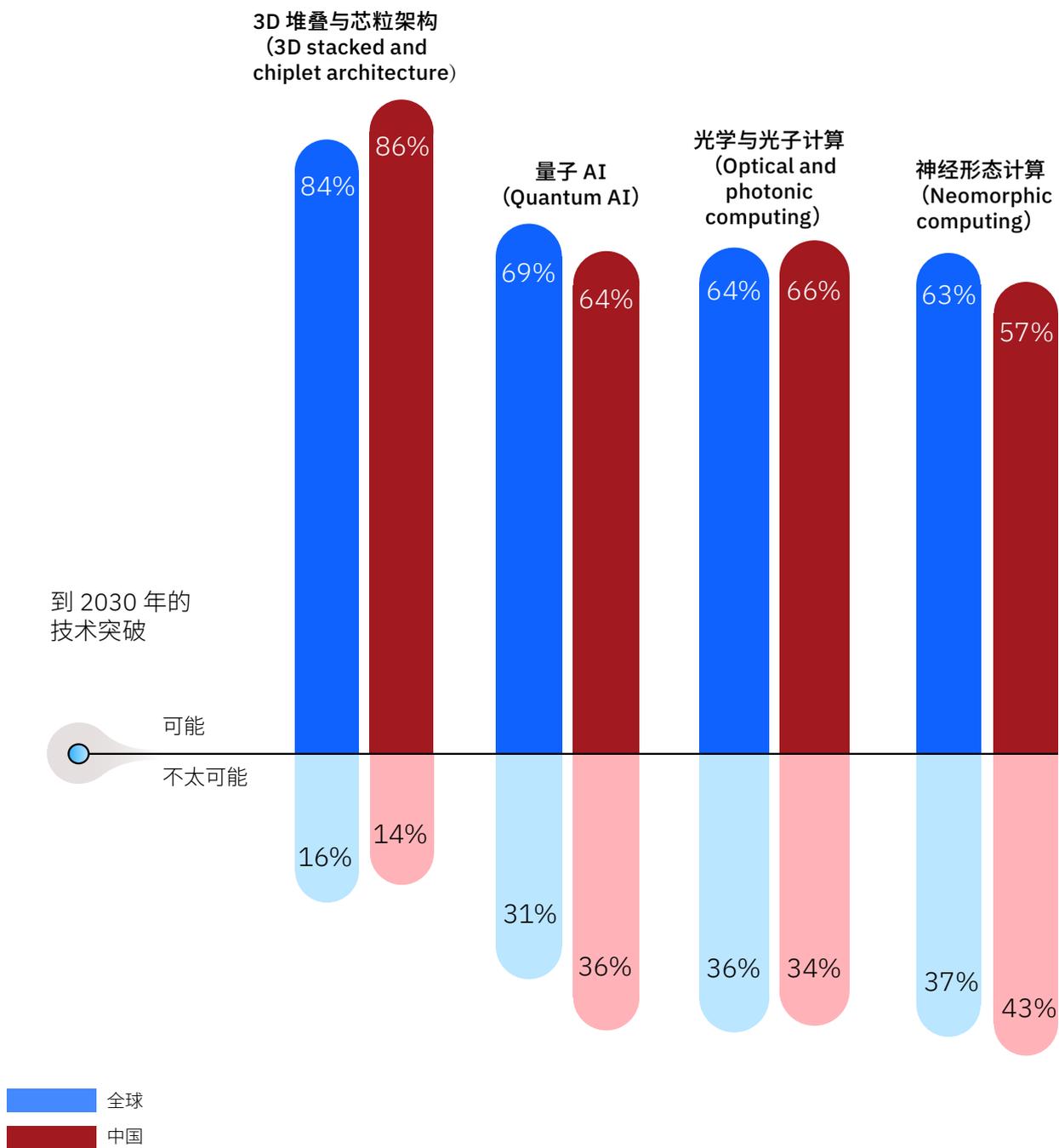
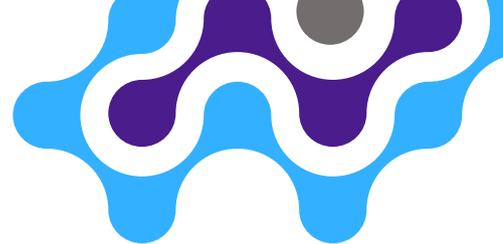


图6

下一代技术有望满足 AI 需求。





行动指南



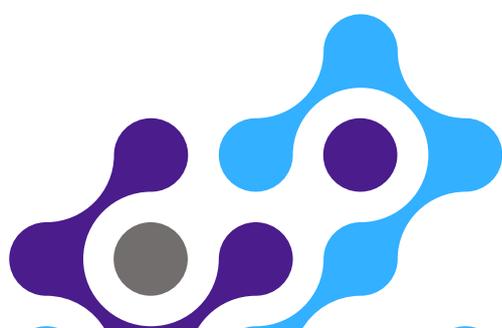
探寻定制化破局点，构筑差异化竞争优势。 识别最能从定制组件中受益的 AI 能力，再以差异化成果为目标，精准评估这些 AI 能力的战略价值。并携手战略合作伙伴，共同探索替代性技术路径，例如先进封装与芯粒架构，以提升 AI 竞争力。

开展小规模试点，前瞻验证方案可行性。 挑选高潜力场景，做早期体验者与反馈者。探索协作新路径，譬如共建仿真平台或合投测试中心，实现成本共担、风险分散与知识速萃。

突破当下思维局限，勾勒未来 AI 应用场景。 预判突破技术的到来，预置资源与流程，缩短落地周期。为员工配备最先进的 AI，将重复性任务交给 AI 代理执行，释放员工创造力。

“现代药物研发，需要先进算力与 AI 的双翼支撑。我们正开辟全新路径，开发定制芯片——采用相同的硬件，但通过专用的 AI 软件和模拟电路进行优化——从而加速药物研发过程。”

Venkat Mattela
Ceremorphic
首席执行官



观点

解决 AI 能源问题

全球正投身于利用海量数据训练 AI 模型的算力竞赛，但这也伴随着巨大的能源代价。AI 训练消耗巨额电力，新建 AI 训练数据中心与配套发电厂比邻而居的情况也日益普遍。国际能源署 (IEA) 预测，至 2030 年，全球数据中心的耗电量预计将增加一倍以上，规模将与日本一整年的用电量相当。⁶

“我们正通过开发专用 AI 芯片与 2 纳米制程，致力于将功耗大幅降低约 75%。”

Atsuyoshi Koike
Rapidus
首席执行官

调研显示，尽管 57% 的数据中心运营商对其芯片供应商在提升能效方面的努力表示满意，但仅有 40% 的供应商对自身的进展感到满意。这一现象或许源于当前 AI 的迫切需求主要聚焦于大模型训练上：对 AI 训练的投资预计在 2028 年左右达到顶峰，此后到 2030 年，行业关注重心将转向 AI 推理。⁷

尽管当前训练 AI 模型的竞争热潮已将能效问题搁置一旁，但该问题依然存在。并且，当 AI 主战场驶向边缘推理与终端部署，能效便是下一个决胜关口。

芯片买卖两端同步出招，各攻其点。例如，83% 的数据中心和高性能计算制造商正在开发针对特定任务优化的专用 AI 处理器；78% 的工业用户已转向轻量化 AI 模型部署。此外，70% 的半导体设备与材料供应商正致力于节能制造工艺，其中 65% 期望 AI 可以赋能提速。

“革新不止于芯片硬件，更在于芯片内运行软件的持续优化与效率提升。而我们，从未停下精益求精的脚步。”

Nvidia
总监

附录

芯片类型与应用指南

AI 的运行依赖于多种半导体各司其职。在训练阶段，为提供输出准确性，AI 需借助高性能的逻辑和内存芯片处理海量数据，并进行数百万乃至数十亿次的参数优化。至推理阶段，当训练完成的模型开始执行任务时，则需依靠专用性更强的芯片来实现性能优化。本术语表可助您依据不同的 AI 需求与行业特定用例，甄选最适宜的芯片类型。

	在 AI 开发和 应用中的作用	主要 用例
逻辑芯片类型		
GPU (图形处理器)	在训练中占主导地位，也用于推理	数据中心
NPU (神经处理器)	主要用于推理	智能手机和边缘设备
ASIC (专用集成电路)	训练和推理，针对特定工作负载	云服务巨头提供的 AI 服务
TPU (张量处理器)	训练和推理	Google Cloud AI
AI SoC (系统级芯片)	主要用于推理	智能手机和边缘设备的全芯片集成
内存芯片类型		
DRAM (动态随机存取存储器)	在训练和推理中起支持作用	为 AI 加速器提供数据
HBM (高带宽内存)	对训练至关重要，对推理起支持作用	训练前沿 AI 模型

作者

中国洞察作者

杨华雨

IBM Consulting 合伙人
中国区电子与电气行业总经理
ericyang@ibm.com

蔡宇明

IBM Consulting 副合伙人
中国区 MES 解决方案交付负责人
nico.cai@ibm.com

王莉

IBM 商业价值研究院
高级咨询经理
gbswangl@cn.ibm.com

全球洞察作者

Rami Ahola 博士

IBM Consulting
工业品制造全球行业负责人
linkedin.com/in/ramiahola/

Pushkar Apte 博士

SEMI
战略技术顾问
linkedin.com/in/pushkar-apte-5363/

Stephen Pierce

IBM Consulting 合伙人
离散制造 4.0 与工业制造负责人
linkedin.com/in/spierceibm/

Noriko Suzuki

IBM 商业价值研究院
电子与汽车行业负责人
linkedin.com/in/norikosuzuki/

特别感谢

Mukesh Khare

IBM 半导体

业务总经理

Noboru Itoh

IBM Technology

日本半导体副总裁

Masahito Terakado

IBM Consulting

日本半导体业务负责人

Norishige Morimoto

IBM Research

日本副总裁

Gaurav Kapoor

IBM Consulting

美国半导体

首席客户合伙人

Michael Kuo

IBM Consulting

中国台湾台积电

首席客户合伙人

Yoshihiro Shimada

IBM Consulting

日本电子行业负责人

Yoshifumi Sakamoto

IBM Research

边缘计算首席工程师

Yuji Tomita

IBM Research

SiView

IBM 商业价值研究院

Sara Aboulhosn, 副创意总监

Angela Finley, 设计主管

Jiu Hanashima, 顾问

Tegan Jones, 编辑主管

Christopher Nowak, 研究主管

Tihomir Trifonov, 视觉设计师

Michael Tucker, 编辑主管

520

半导体用户企业

12

国家

280

半导体供应企业

14%
美国

5%
英国

5% 荷兰
9% 德国

5% 3% 瑞士
法国

12% 中国

12% 日本

11% 韩国

8% 印度

12% 中国台湾

4% 新加坡

研究方法

IBM 商业价值研究院 (IBM IBV) 与 Phronesis Partners 携手合作，于 2025 年第二季度，对来自 12 个国家的 520 家半导体用户企业（“芯片买家”）及 280 家半导体供应企业（“芯片供应商”）的最高管理层开展了调研，其中有 100 家来自于中国大陆。

用户企业覆盖九个在其产品和服务中采用先进半导体的行业领域——汽车、数据中心、消费电子、工业机械、电信运营商、医疗设备、智能手机、PC/ 平板 / 游戏设备，以及 HPC/ 大型机。半导体供应企业则涵盖九个主要环节——无晶圆厂、晶圆代工厂、材料、机械设备、集成器件制造商 (IDM)、先进封装、分销商、以及电子设计自动化 / 知识产权 (EDA/IP) 厂商。

调研设计包含多种问题形式，包括多项选择、数值题、李克特量表问题，聚焦企业对 AI 技术进展的期望、实际成果与未来展望，以及整合了 AI 技术的产品与半导体战略的实践情况。

对这一全面数据集的分析需要一个复杂的多维分析框架。本研究采用了多种互补性分析技术，每种技术均旨在揭示数据结构中的独特洞察。

首先，通过组级比较分析系统评估供应侧与需求侧因素，确定它们在推动 AI 发展路径中的相对重要性。为捕捉数据中复杂的非线性关系，研究团队部署了多层感知器 (MLP) 神经网络和序数逻辑回归，以识别主导本地化采购驱动机制的复杂模式。

随后，应用 K 均值聚类算法，根据各组织的 AI 成熟度水平和组织韧性特征，将其系统划分为同质细分群体。同时，对研发预算分配进行了细分分析，以识别不同组织间的独特投资模式。

这种集成分析方法确保了对数据集的全面检视，同时在整个研究过程中保持了方法的严谨性。

此外，研究团队还对行业高管进行了 14 场深度访谈，以获取更细致入微的战略洞见，并将这些定性的洞察也纳入了分析中。

IBM 商业价值研究院

IBM 商业价值研究院 (IBM IBV) 自创建以来持续提供有研究支持和技术支持的战略洞察，帮助领导者做出更明智的业务决策。凭借 IBM 在商业、技术和社会交叉领域的独特地位，我们每年都会针对成千上万高管、消费者和专家展开调研、访谈和互动，将他们的观点综合成可信赖的、振奋人心和切实可行的洞察。需要 IBV 最新研究成果，请在 [ibm.com/ibv](https://www.ibm.com/ibv) 上注册以接收 IBV 的电子邮件通讯。访问 IBM 商业价值研究院中国官网，免费下载中文研究报告：<https://www.ibm.com/ibv/cn>。

选对合作伙伴，驾驭多变的世界

在 IBM，我们积极与客户协作，运用业务洞察和先进的研究方法与技术，帮助他们在瞬息万变的商业环境中保持独特的竞争优势。

关于研究洞察

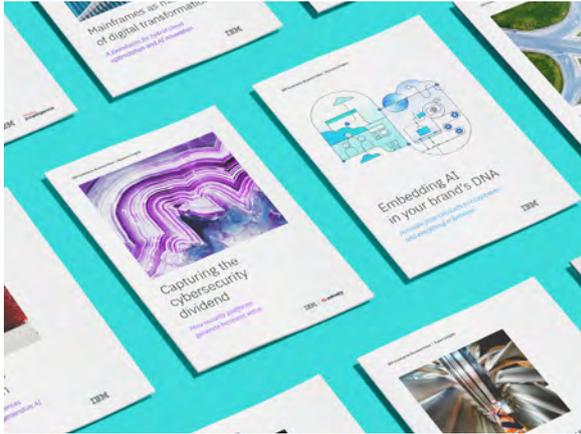
研究洞察致力于为业务主管就公共和私营领域的关键问题提供基于事实的战略洞察。洞察根据对自身主要研究调查的分析结果得出。要了解更多信息，请联系 IBM 商业价值研究院：ibv@us.ibm.com。

关于 SEMI

SEMI® 是全球半导体与电子设计制造供应链的行业协会，连接了全球 3,000 多家会员公司和 150 万专业人士。通过政策倡导、人才培养、可持续发展、供应链管理等多个项目，加速会员合作，共同应对行业重大挑战。SEMICON® 展会与活动、技术社区、标准制定以及市场情报，助力会员在设计、器件、设备、材料、服务及软件领域实现业务增长与创新，推动电子产业更智能、更快速、更安全的发展。了解更多信息，请访问 www.semi.org。

备注和参考资料

- 1 “历史新高！SEMI 预计 2025 年全球半导体设备总销售额达 1255 亿美元”。证券时报网 . 20250723. <https://www.stcn.com/article/detail/2725900.html>
- 2 缪琦 . “SEMI：全球半导体市场今年仍将两位数增长，三大趋势影响未来”。第一财经 . 20250306. <https://www.yicai.com/news/102500385.html>
- 3 Jones, Cristela. “Samsung’s massive Texas plant scores \$16B AI-chip deal from Elon Musk.” MSN. July 28, 2025. <https://www.msn.com/en-us/news/technology/samsungs-massive-texas-plant-scores-16b-ai-chip-deal-from-elon-musk/ar-AA1JrJJw?apiversion=v2&noservercache=1&domshim=1&renderwebcomponents=1&wcseo=1&batchservertelemetry=1&noservertelemetry=1-telcs>
- 4 “Apple says Samsung will supply chips from Texas factory.” Reuters. August 7, 2025. <https://www.reuters.com/business/apple-says-samsung-will-supply-chips-texas-factory-2025-08-06/>
- 5 “Fab Runs On Code.” TSMC IT Careers. Accessed September 4, 2025. https://www.tsmc.com/static/english/careers/it_career/index.html
- 6 “AI is set to drive surging electricity demand from data centres while offering the potential to transform how the energy sector works.” International Energy Agency. April 10, 2025. https://www.iea.org/news/ai-is-set-to-drive-surging-electricity-demand-from-data-centres-while-offering-the-potential-to-transform-how-the-energy-sector-works?utm_source=chatgpt.com
- 7 Walters, Asya, Andy Walker, and Moe Kelley. “Rethinking AI Demand Part 1: AI Data Centers Are Experiencing a Surge of Training Demand - What Happens When the Surge is Over?” A&M. February 25, 2025. https://www.alvarezandmarsal.com/insights/rethinking-ai-demand-part-1-ai-data-centers-are-experiencing-surge-training-demand-what?utm_source=chatgpt.com



订阅 IdeaWatch 电子月刊

前沿洞察，触手可及。

由 IBM 商业价值研究院倾力呈现。IBM 商业价值研究院连续两年荣膺 Source Global Research “全球思想领导力排行”榜首。

基于深度研究的战略洞见、数据与分析，助您优化商业决策，智胜科技投资。

立即订阅: ibm.co/ideawatch



© Copyright IBM Corporation 2025

国际商业机器(中国)有限公司
北京市朝阳区金和东路 20 号院 3 号楼
正大中心南塔 12 层
邮编:100020

美国出品 | 2025 年 11 月

IBM、IBM 徽标、ibm.com、IBM Z、IBM z16 和 watsonx 是 International Business Machines Corporation 在世界各地司法辖区的注册商标。其他产品和服务名称可能是 IBM 或其他公司的商标。以下 Web 站点上的“Copyright and trademark information”部分中包含了 IBM 商标的最新列表：ibm.com/legal/copytrade.shtml。

本档为自最初公布日期起的最新版本，IBM 可能随时对其进行更改。IBM 并不一定在开展业务的所有国家或地区提供所有产品或服务。

本文档内的信息“按现状”提供，不附有任何种类的（无论是明示的还是默示的）保证，包括不附有关于适销性、适用于某种特定用途的任何保证以及非侵权的任何保证或条件。IBM 产品根据其提供时所依据的协议条款和条件获得保证。

本报告的目的仅为提供通用指南。它并不旨在代替详尽的研究或专业判断依据。由于使用本出版物对任何企业或个人所造成的损失，IBM 概不负责。

本报告中使用的数据可能源自第三方，IBM 并未对其进行独立核实、验证或审查。此类数据的使用结果均为“按现状”提供，IBM 不作出任何明示或默示的声明或保证。

扫码关注 IBM 商业价值研究院



官网



微博



微信公众号

