

# 芯片“智”造：半导体企业正利用人工智能设计更优芯片，速度更快、成本更低且效率更高

近期机器学习的发展正赋能芯片企业解决芯片设计史上的一大难题：如何在平方英寸的空间排列上千亿个晶体管？

Jeff Loucks, Duncan Stewart, Christie Simons, Brandon Kulik, Roger Chung, Mark Lian, Leo Chen

人工智能正在迅速成为人类芯片工程师的强大助手，助力高效完成极度复杂的半导体设计工作。德勤全球预测，2023年全球半导体企业将投入3亿美元，以利用内部自有或第三方人工智能工具开展芯片设计，<sup>1</sup>且未来四年这一数字将每年增长20%，到2026年将超过5亿美元。<sup>2</sup>2023年全球半导体市场规模预计将达到6,600亿美元，<sup>3</sup>相比之下这一数额虽然不大，但对于投资所

带来的超常回报而言却意义非凡。人工智能设计工具正在赋能芯片制造商拓宽摩尔定律的边界，节约时间和资金成本，缓解人才短缺困境，甚至带动过时芯片设计迈入现代化新时代。与此同时，这些工具亦能增强供应链安全保障，助力延缓下一波芯片短缺潮。换言之，用于芯片设计的人工智能软件工具的单用户授权费用虽可能需要数万美元，但利用这些工具设计的芯片价值可达数十亿美元。

## 与时间赛跑：先进人工智能极大加速芯片设计进程

数十年来，电子设计自动化 (EDA) 供应商已针对芯片设计开发了多种工具——2022年这一行业规模达到100亿美元以上，且正以每年8%的速度增长。<sup>4</sup>EDA工具通常采用基于规则的系统 and 物理仿真，以协助人类工程师开展芯片设计与验证。部分工具甚至融合了基础的人工智能。然而过去一年，在芯片制造商和科技企业已开发出自有人工智能设计工具的同时，规模最大的EDA公司也已开始推出先进的人工智能赋能型工具。<sup>5</sup>这些先进工具并不仅仅是试验品，而是正被应用于诸多真实芯片设计场景中，每年创造的价值可达数十亿美元。尽管并不能取代人类设计师，但这些工具在速度和成本效益方面具有显著的互补性优势，大大增强了芯片制造商的设计能力。

芯片设计与制造过程极为复杂。先进人工智能可在以下三个主要方面提供助力：

**制造全新及更优芯片：**采用10nm以下制程节点的芯片普遍应用于智能手机、计算机和数据中心，在芯片市场中增长最快，<sup>6</sup>也是目前最具盈利能力的部分。但其制造成本亦是最高，单个全新芯片设计投入高达5亿美元。<sup>7</sup>相比传统方法，先进人工智能工具可显著提高这些芯片的设计速度，从而降低成本。

**提升过时芯片性能：**2022年全年售出的芯片中，有三分之二采用65nm或更大的制程节点，而这是几十年前的过时技术。<sup>8</sup>将这些过时的芯片设计转至更为先进的制程节点（即“缩小”），能让这些芯片实体更加小巧、能效更高，且并不依赖于过时淘汰的制造设备。借助先进人工智能工具，芯片制造商能更快速地以更低成本完成“缩小”过程。

**弥补芯片人才短缺：**2022年全球芯片行业从业人数约为200万人，然而随着美国、欧盟和中国日益推进芯片自主化战略，到2030年芯片行业还需增加100万人。<sup>9</sup>先进人工智能工具作为一个可弥补人才短缺的方法，其重要性将会日益增长。

图 1

## 先进人工智能芯片设计工具将迅速增长，预计将为EDA工具的两倍以上、芯片销量增长的三倍以上

芯片销量、EDA工具及先进人工智能设计工具五年年均复合增长率预测（2023–2028）



资料来源：WSTS；Global Market Insights；及德勤全球。

芯片设计需要经过三个主要阶段：系统级设计、寄存器传输级设计（RTL）以及最终的物理电路设计。先进人工智能工具可真正发挥作用的，便是在最后的电路设计阶段。

芯片设计的目标是优化三个变量指标——功耗、性能和面积（PPA），从而使制造出来的芯片具有最小用电量、最大处理速度以及最小尺寸。采用传统工具进行PPA优化不仅耗时漫长且需要大量人力：设计迭代可能耗时长达数周，且这些迭代对PPA的提升往往极为有限。而从芯片设计到实体落地，再到设计与实体落地的评估、测试和仿真，可需要长达数年的时间。

芯片内部含有数以十亿计的晶体管，表现为不同的组合模块（包含存储器子系统、计算单元、控制逻辑系统及电源）和标准单元。在高度复杂的芯片中，这些组合模块被最长达50千米的线路连接在一起。若组合模块未实现最优化排列，其连接所需的线路和空间则会更长、更大。各组件之间产生的非预期电荷量（称为寄生效应）可降低芯片的性能，增加功耗。

先进人工智能工具可用于检测人类完成的设计，发现导致功耗增加、性能下降或空间利用效率低的布局错误，提出改进建议，并对此进行仿真验证和检测。这些工具可从过往迭代中学习提升，以最大限度地改善PPA。但是真正革命性的一点，是先进人工智能可以自主完成这一过程，相比利用传统EDA工具的人类设计师，能够实现更优的PPA表现，某些情况下甚至能在几个小时内完成，仅需一名设计工

程师协助，而无需整个工程团队工作数周乃至数月的时间。

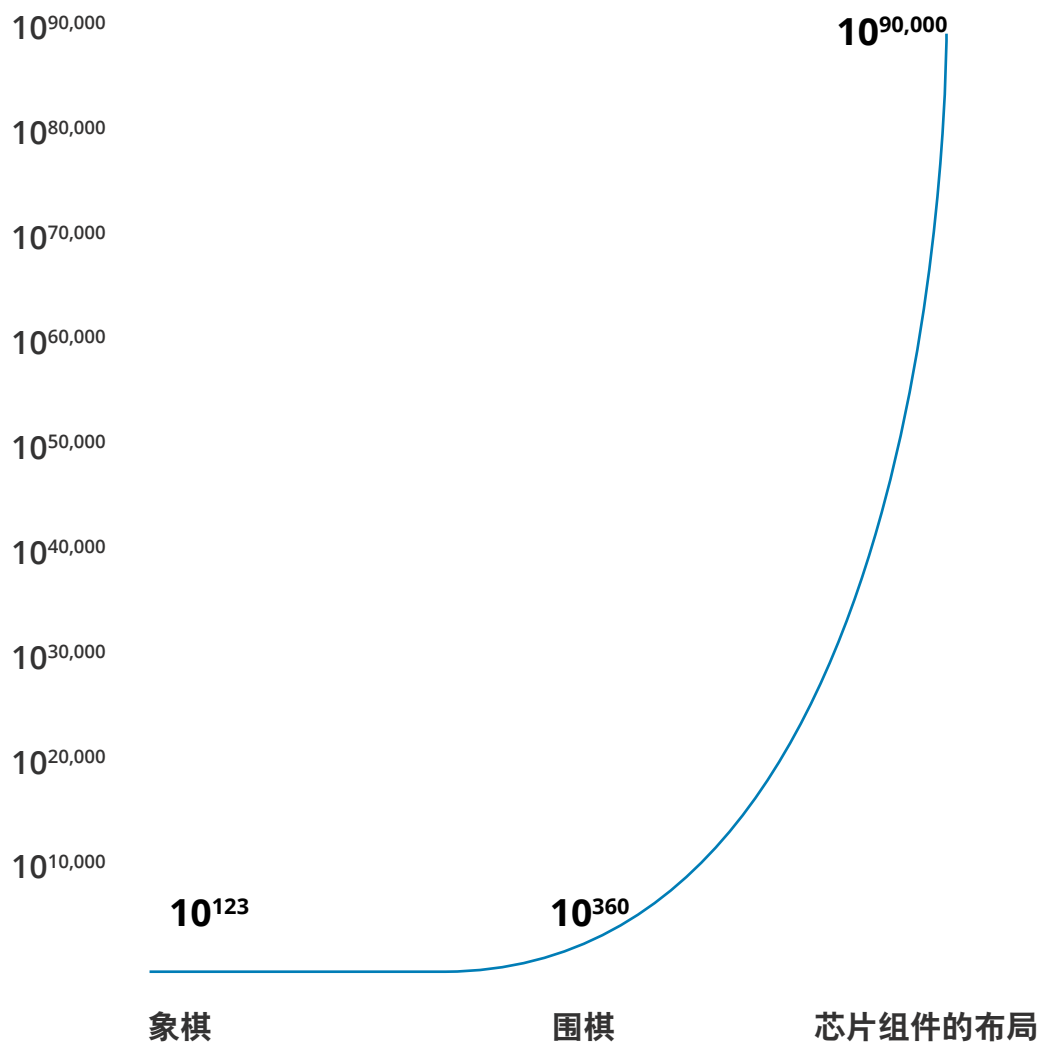
这些先进的人工智能能力基本可以分为两种类型：图神经网络（GNN）和强化学习（RL）。图神经网络是专门用于分析图形（包含“节点”和“边缘”的数据结构，节点可以是任何对象，而边缘则定义了节点之间的关系）的一种机器学习算法。<sup>10</sup>传统的深度学习神经网络在图形分析方面较为棘手，<sup>11</sup>而图神经网络则能够从图形中提取信息，有效预测图形之间的联系，并在保留关键联系的同时将节点重新排列。<sup>12</sup>由于芯片的结果本质上与图形类似——宏模块和标准单元相当于节点，将其连接在一起的线路则相当于边缘——图神经网络是进行芯片分析和优化的绝佳之选。

强化学习可将物理芯片设计变成一个图形优化“游戏”。这与谷歌在战略棋盘游戏——围棋上用于击败人类冠军的是同一种技术。围棋比象棋甚至更为复杂，且曾被认为是超出人工智能能力范围的游戏。物理芯片设计的复杂程度远远高于围棋（图2），但强化学习依然可以同样的方式应对。它利用数以千计的“棋局”——芯片布局规划进行训练，通过仿真芯片设计找到最优的PPA排列。人工智能生成布局规划后，人类设计师可对有效优化PPA的设计（如减少了线路长度、拥挤度、密度、功耗和面积等<sup>13</sup>）给予“奖励”，并对不理想的设计给予“处罚”，从而进一步强化布局规划设计。久而久之，这种持续不断的强化能够提升强化学习系统，使其学习到如何自主生成更优设计。<sup>14</sup>

图 2

## 人工智能在象棋和围棋上击败人类已令人大开眼界，若是设计芯片呢？

相比象棋和围棋，芯片设计在配置上拥有的可能性远远更多



资料来源：摘自新思科技“什么是设计空间优化？”，2020年7月21日。

采用神经网络和强化学习结合实现的PPA，其表现与由经验丰富的设计师制造出的PPA不相上下，甚至更优，且所需要的人类工程师人数更少，耗时亦远远更短。这在现实中已有不少案例，近期有：

- 麻省理工借助人工智能工具开发电路设计，其能量效率是人类设计的电路的2.3倍。<sup>15</sup>
- 联发科采用人工智能工具调整一个关键处理器部件，使其尺寸减小了5%，能耗降低了6%。<sup>16</sup>
- Cadence成功将一枚5nm移动芯片的性能提升了14%，能耗减少了3%，仅需人工智能工具加上一名工程师用10天的时间完成，而无需10名工程师耗费数月之久。<sup>17</sup>
- 谷歌母公司Alphabet持续稳定产出芯片布局规划，PPA指标均优于经验丰富的人类工程师，且耗时在六小时以内，而非数周乃至数月。<sup>18</sup>
- 英伟达利用自有强化学习工具进行电路设计，在确保具备相同性能的同时，成品尺寸较人类工程师采用EDA工具设计的电路降低25%。<sup>19</sup>

## 中国：AI技术助力EDA发展

芯片设计环节繁多，精细且复杂，EDA工具对于提升芯片设计效率，优化芯片设计，保证芯片功能发挥着极为重要的作用。目前，美国三大EDA公司(Synopsys、Cadence、Mentor) 占据全球EDA市场超过60%的市场份额<sup>1</sup>，绝大部分芯片设计公司都需要三巨头的EDA工具。国产化率仅10%左右的国产EDA<sup>2</sup>，面对发展了二三十年的EDA三巨头的技术和商业壁垒，想要进一步提升国产化率面临极大挑战。目前，中国的AI技术全球领先，这种优势在一定程度上有助于短期内加快追赶步伐。还有不可忽略的云技术，随着云计算的发展，云上设计芯片能够减少芯片设计流程中耗时较多的芯片设计验证时间。受益于半导体产业向中国转移趋势，中国EDA市场将以14.71%的CAGR增长，预计在2025年将达到27.4亿美元的市场规模。

### 小结

主要芯片制造商和设计师如今正采用先进人工智能设计芯片，甚至在先进节点也不例外。事实上，部分芯片的复杂性日益增长，先进人工智能的使用可能很快便成为必然要求：例如，新思科技的最新芯片设计包含超过1.2万亿个晶体管 and 40万个人工智能优化核心。<sup>20</sup>

借助基于云的EDA服务，先进人工智能亦正在加速普及，潜在市场进一步扩大。成功上云后，技术水平和计算能力不足的小型企业也能够充分获取先进人工智能能力，不再仅限于少数行业专家和市场领导者。<sup>21</sup>

规模最大的半导体公司甚至可以利用先进人工智能开发全新服务，创造新的收入来源。通过进一步扩大自身的神经网络和强化学习能力，这些公司不仅能够自主开展芯片设计，同时亦可为其重要客户提供芯片设计或联合设计服务，包括联合开发垂直专用芯片。

对芯片行业而言，人工智能的作用可并不仅仅局限于芯片设计。例如，人工智能可用于晶圆的外表检查，其检查次数可达到近九次，从而提升故障检测能力。<sup>22</sup>借助人工智能，芯片公司还能有效应对供应链方面的挑战，如管理半导体组装和测试外包提供商网络。<sup>23</sup>

近几年来，我们曾针对人工智而设计专门的芯片；如今，由人工智能设计芯片已成必然之势。未来将会如何发展？人工智能有望与人类联手，共同设计驱动其自身运行的硬件和软件，开启21世纪新一轮创新浪潮！

## 尾注

1. 德勤全球预测，2022年主要供应商的第三方人工智能芯片设计软件市场规模约为1.5亿美元，2023年将增长至超过2亿美元。此外，我们预测大型芯片公司内部自用的人工智能设计工具亦具有相当规模的价值。
2. 德勤全球根据EDA公司公开声明和相关分析报告预测的增长率。
3. World Semiconductor Trade Statistics, "The World Semiconductor Trade Statistics (WSTS) has released its new semiconductor market forecast generated in August 2022," press release, August 22, 2022.
4. Global Market Insights, *Electronic design automation market report*, 2020.
5. 在EDA工具中融合机器学习元素已有数十年时间，而图神经网络和强化学习等先进人工智能技术的使用则是新兴趋势，并已经显著提升人工智能在芯片设计方面的成效。
6. John Ciacchella et al., *2022 semiconductor industry outlook*, Deloitte, 2022.
7. International Business Strategies (IBS), 2021.
8. Ciacchella et al., *2022 semiconductor industry outlook*.
9. 德勤全球采用自上而下（国家/地区最新报告的直接就业数据）和自下而上（所有大型公司报告的员工数据）两种方式估算2021年全球半导体行业的直接就业人数。鉴于到2030年该行业收入规模将增长80%，而行业集中程度相较于当前将会降低（因而单位营收则需要更多员工），我们假定行业将需要增加约50%的员工。
10. Abid Ali Awan, "A comprehensive introduction to graph neural networks (GNNs)," DataCamp, July 2022.
11. 如欲直观地了解关于为何神经网络难以有效分析图形数据，以及为何图神经网络表现更好的原因，参见Ben Dickson, "What are graph neural networks (GNN)?"，VentureBeat, 2021年10月13日；如欲从技术角度了解更多，参见DataCamp, "A comprehensive introduction to graph neural networks (GNNs)"，2022年7月。
12. Dickson, "What are graph neural networks (GNN)?"
13. Ed Targett, "AI outperforms humans in chip design breakthrough," The Stack, June 10, 2021.
14. BBC News, "Go master quits because AI 'cannot be defeated'," November 27, 2019.
15. Will Knight, "Need to fit billions of transistors on a chip? Let AI do it," *Wired*, July 9, 2021.
16. James Morra, "Cadence taps AI technology to speed up system design," *Electronic Design*, June 13, 2022.
17. John Koon, "Improving PPA with AI," *Semiconductor Engineering*, May 12, 2022.
18. Azalia Mirhoseini et al., "A graph placement methodology for fast chip design," *Nature* 594 (2021): pp. 207-12.
19. Rajarshi Roy, Jonathan Raiman, and Saad Godil, "Designing arithmetic circuits with deep reinforcement learning," NVIDIA Developer, July 8, 2022.
20. Stelios Diamantidis, "Why now is the time to create an AI strategy for chip design," Synopsys blog, June 16, 2021.
21. Jeff Loucks, *Artificial intelligence: From expert-only to everywhere*, Deloitte Insights, December 11, 2018.
22. Tobias Schlosser et al., "Improving automated visual fault inspection for semiconductor manufacturing using a hybrid multistage system of deep neural networks," *Journal of Intelligent Manufacturing* 33 (2022): pp. 1099-1123.
23. Deloitte, "Supply chain and network operations: Enterprise technology and performance," accessed October 26, 2022.

## 关于作者

### **Jeff Loucks | [jloucks@deloitte.com](mailto:jloucks@deloitte.com)**

Jeff Loucks is the executive director of Deloitte’s Center for Technology, Media & Telecommunications, Deloitte Services LP. He conducts research and writes on topics that help companies capitalize on technological change.

### **Duncan Stewart | [dunstewart@deloitte.ca](mailto:dunstewart@deloitte.ca)**

Duncan Stewart is the director of TMT Research for Deloitte Canada and is a globally recognized specialist on the forecasting of consumer and enterprise technology, media & telecommunications trends. He presents regularly at conferences and to companies on marketing, technology, consumer trends, and the longer-term TMT outlook.

### **Christie Simons | [csimons@deloitte.com](mailto:csimons@deloitte.com)**

Christie Simons is the Audit & Assurance Technology, Media & Telecommunications (TMT) industry leader with Deloitte & Touche LLP. She specializes in serving fast-growing, dynamic venture-backed companies and large complex publicly listed technology clients with external audit, IPO, M&A, and SOX services.

### **Brandon Kulik | [bkulik@deloitte.com](mailto:bkulik@deloitte.com)**

Brandon Kulik leads the semiconductor industry segment within Deloitte Consulting LLP’s Technology, Media & Telecom industry practice. He has more than 20 years of experience helping technology companies with market strategy, operational, talent, and technology solutions.

## 致谢

The authors would like to thank **Ariane Bucaille, Gillian Crossan, Dan Hamling, and Karthik Ramachandran** for their contributions to this chapter.