

News & Highlights

精简指令集架构加速芯片研发

Chris Palmer

Senior Technology Writer

此前被人们循序渐进采用的流水线式计算机芯片指令集架构 (ISA) —— 第五代精简指令集计算机 (RISC-V), 现已正式进入高速发展阶段。2022年2月初, 芯片巨头英特尔 (美国加利福尼亚州圣克拉拉) 宣布了对 RISC-V 生态系统进行大笔投资, 这是英特尔所发起的 10 亿美元基金的一部分; 该项基金用于支持英特尔与其他芯片初创公司之间的合作[1-2]。基金的设立不但支持了英特尔发展其代工业务的计划, 还与该公司同时宣布加入 RISC-V International 组织的举动相得益彰。该组织总部位于瑞士苏黎世, 负责管理 RISC-V (RISC-V 是 2010 年首次引入的免费开放式标准 ISA) 的内核和新扩展编目[1-2]。

RISC-V International 首席执行官 Calista Redmond 表示: “英特尔与几家 RISC-V 成员的合作, 向业界发出了一个巨大的信号, 即人们对 RISC-V 的兴趣与采用正与日俱增。认识到 RISC-V 将是其代工事业成功的一个重要组成部分后, 为更得心应手地设计 RISC-V 解决方案, 英特尔正对其大力投资。”

目前, RISC-V 市场虽然仍相对较小, 但预计将会快速增长。咨询和分析公司德勤全球 (美国纽约市) 预测, 基于 RISC-V 的处理内核的市场将在 2022 年比 2021 年翻一番, 并在 2023 年再次翻一番[3]。不过, RISC-V 的收益增长速度会更加缓慢, 从 2021 年的不到 4 亿美元增加到 2023 年的近 8 亿美元, 再到 2024 年接近 10 亿美元[3]。

在基本层面上, 芯片中的 ISA 可让软件直接控制硬件。现在驱动世界上大多数设备的芯片均使用了两种成熟

的专有 ISA。在笔记本电脑、台式电脑和服务器硬件芯片中, 英特尔 x86 在过去几十年一直保持着主导地位。但由于其耗电量大, 相比之下, 大多数智能手机则采用了基于 ARM 的芯片。现在, 美国、欧洲, 尤其是亚洲越来越多的公司, 均将 RISC-V 视为对这些根深蒂固标准的一种极具吸引力的替代和补充选项。

RISC-V 是多年来创建的众多基于 RISC 的 ISA 之一 (包括 ARM 在内)。2022 年 3 月, 有四人因此荣获了 2022 年查尔斯·斯塔克·德拉普尔奖 (Charles Stark Draper Prize) [4]。将它称为“精简指令集计算机”也反映了获奖者们以及他们众多合作者的目标, 即通过降低核心可执行指令集的复杂性来构建更小、更低功耗但性能更好的处理器。四位获奖者分别是英国曼彻斯特大学计算机工程教授 Stephen B. Furber、美国加利福尼亚州帕洛阿尔托斯坦福大学计算机科学与电子工程教授 John L. Hennessy、美国加利福尼亚大学伯克利分校名誉教授 David A. Patterson 以及英国剑桥大学计算机科学家 Sophie M. Wilson [4]。与 x86 的 1000 多条指令和 ARM 的大约 500 条指令相比, 采用了 RISC-V 的内核只有 47 条指令, 简直是微不足道[2, 5]。而且, RISC-V 还是一种可扩展的 ISA, 这意味着各公司不但可实现指令集内核, 还可通过添加已有定义或自定义的指令来扩展内核, 从而创建自己的定制处理器。

此特性让基于 RISC-V 的芯片设计易于修改, 在应用于广泛的数字应用 (从高性能计算、自动驾驶车辆和物联网到存储、图形、人工智能和机器学习等) 时, 具有更大

的灵活性[3]。RISC-V 处理器开发公司 Ventana Micro Systems（美国加利福尼亚州库比蒂诺，英特尔的新合作伙伴）创始人兼首席执行官 Balaji Baktha 表示：“RISC-V 非常先进。由于不受 30 年累积常识的影响，它成为了一个‘全新’的架构。”

除了提供灵活性外，开放式标准也使 RISC-V 成为一个有吸引力的选项，特别是对于那些既希望设计定制芯片，又不希望花费数百万美元许可费来使用专有芯片架构的初创公司而言。这两个因素让 RISC-V 成为了诸如计算存储巨头希捷（美国加利福尼亚州库比蒂诺）、西部数据（美国加利福尼亚州圣何塞）以及电子商务巨头阿里巴巴（中国杭州）等公司间交口称誉的芯片架构[6]。

作为一个开放式标准，对于那些可能因出口限制而失去或担心失去 x86 或 ARM 授权的公司而言，RISC-V 极具吸引力[7]。因为选择 RISC-V，就可以消除他们所担心的这个问题。RISC-V International 组织三分之二的成员是中国公司，其中包括华为（中国深圳）和阿里云，而且多家中国大公司也已宣布打算使用 RISC-V 芯片[6]。此外，中国在 2021 年年底就成立了一个专门从事 RISC-V 处理器项目研究的机构[8]，阿里巴巴宣布将致力于把安卓系统移植到 RISC-V [9]，中国的两家公司 Deep Computing 和 Xcalibyte 也已经开始在深圳制造世界上第一台基于 RISC-V 的笔记本电脑[10]。

通过加入 RISC-V 的行列，英特尔旨在帮助以 RISC-V 开发为重点的公司，包括已宣布的合作伙伴晶心科技（中国台湾新竹）、Esperanto Technologies（美国加利福尼亚州山景城）、SiFive（美国加利福尼亚州圣马特奥）和 Baktha 创办的 Ventana Micro Systems 公司，希望通过为他们优先安排生产流程，并与他们合作开发优化设计的制造技术，来更快地进行创新[1-2]。

英特尔设立 10 亿美元基金的主要目标是实现一种具有成本效益的模块化设计，将多种类型的处理器内核集成在单个芯片封装中[1-2]。这将有助于这家芯片巨头的业务英特尔代工服务（Intel Foundry Services）——制造包含多个小芯片裸片的芯片，并定制 x86、ARM 和 RISC-V 内核的组合品。Baktha 说：“人们总是问英特尔是否将 RISC-V 视为一种威胁。其实，英特尔的投资就给出了答案：当然不会。通过让 RISC-V 与 x86 基础设施并驾齐驱，人们才能够以前不可能达到的水平来实现更多目标，如数据中心创新。”

人们对 RISC-V 的兴趣很大程度上缘于对人工智能应用的需求。Baktha 说，基于边缘的人工智能直接在笔记本电脑等本地设备上运行相关进程和生成数据，而不是在云

中，因此，“计算和人工智能需要集成在一个统一的架构中。RISC-V 是实现这一目标的唯一途径。” 半导体咨询公司 Semico Research（美国亚利桑那州凤凰城）估计，从现在到 2027 年，基于 RISC-V 的人工智能片上系统（SoC，由计算机的所有或大部分组件组成的集成电路）将以每年 74% 的速度增长[11]。此外，Semico 还预测用于消费物联网解决方案的 RISC-V 人工智能片上系统将在同一时间段内每年增长 140%，而面向数据中心服务器等高性能计算机市场的 RISC-V 人工智能片上系统将以每年 98% 的速度增长[11]。

在 2010 年首次发布之时，RISC-V 缺乏机器学习数学（包括矩阵乘法）所需的矢量指令。为了克服这一不足，Esperanto Technologies 的工程师开发了他们自己的指令，并将其集成到新芯片 ET-SoC-1（图 1）中，之后该公司将 ET-SoC-1 作为第一个高性能 RISC-V 人工智能芯片进行销售[2]。ET-SoC-1 是一款专为提高推理运算效率而设计的芯片，旨在人工智能推荐系统（根据人口统计或先前行为等数据向用户提供相关建议）中与强大的图形处理单元竞争，主要面向电子商务和社交媒体公司[12]。ET-SoC-1 的尺寸是其主要竞争对手英伟达（Nvidia）的 A100 芯片的一半，并且其功率仅为 A100 的十分之一[2]。包括三星在内的几家公司目前正在探索在其设备中使用 ET-SoC-1 [13]。

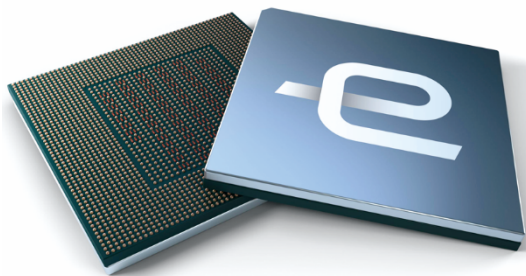


图 1. Esperanto Technologies 的 ET-SoC-1 芯片将 1000 多个 RISC-V 内核封装在一块 570 mm² 的硅片上，运行时仅消耗 20 W。来源：Esperanto Technologies（公共领域）。

其他最近开发出人工智能处理器的公司，也正在将 RISC-V 和定制的机器学习加速技术进行组合。例如，总部位于美国圣何塞的 Ceremorphic 公司于 2022 年 1 月发布的 Hierarchical Learning Processor，就结合了 RISC-V 和 ARM 内核及其自己定制的机器学习和浮点运算部件[14]。英特尔即将推出一款名为 Mobileye EyeQ Ultra 旨在为近乎完全自动驾驶提供智能的芯片，将在一个芯片中集成 12 个 RISC-V 内核，并辅以专有的神经网络加速器，以及一个 ARM CPU [15]。同样转向采用 RISC-V 来支持自动

驾驶的，还有嵌入式人工智能处理器公司 Kneron（美国加利福尼亚州圣地亚哥）。2021年11月，该公司开始批量生产其首个基于新型神经网络的 RISC-V KL530 芯片，用于执行图像分类任务[16]。

2022年6月，已经推出了基于 RISC-V 的 Nios V 处理器的英特尔公司宣布与西班牙巴塞罗那超级计算中心（Barcelona Supercomputing Center）合作建设一个4亿美元的设施，来专门开发用于高性能计算的 RISC-V 处理器，以及用于人工智能和自动驾驶汽车的专用芯片。此次合作的一个主要目标是创建比当今最快的超级计算机强大1000倍的 zettascale 系统，英特尔表示将在未来五年内实现这一里程碑[17–18]。

虽然 RISC-V 的前景看起来很好，但在市场能够完全接受该 ISA 之前，开发者仍然面临着一些挑战[3]。具体而言，某些类型的基于 RISC-V 的新型软件编译器、处理器和计算机设计必须从头开始建立，这是一项耗时而昂贵的工作。此外，虽然 RISC-V 得到了几个 Linux 发行版的支持，包括 Ubuntu、Debian、FreeBSD、NetBSD 和 OpenBSD，但它还没能适应其他一些广泛使用的操作系统。

无论如何，像众多推广者一样，Redmond 关注的是 RISC-V 给计算世界带来的灵活性。Redmond 说：“几十年来，只有拥有庞大设计团队和预算的公司才能实现定制硅。RISC-V 让每个领域和行业实现了设计自由，使各种规模的公司都有可能推动创新发展，并有机会参与竞争。”

References

- [1] Intel launches \$1 billion fund to build a foundry innovation ecosystem [Internet]. Santa Clara: Intel; 2022 Feb 7 [cited 2022 Jul 13]. Available from: <https://www.intel.com/news-events/press-releases/detail/1525/intel-launches-1-billion-fund-to-build-a-foundry>.
- [2] Moore SK. RISC-V AI chips will be everywhere [Internet]. New York City: IEEE Spectrum; 2022 Feb 24 [cited 2022 Jul 5]. Available from: <https://spectrum.ieee.org/risc-v-ai>.
- [3] Stewart D, Bucaille A, Kodama E, Crossan G. RISC-y business: could open chip standard RISC-V gain traction against dominant incumbents? [Internet]. New York City: Deloitte; 2021 Dec 1 [cited 2022 Jul 5]. Available from: <https://www2.deloitte.com/us/en/insights/industry/technology/technology-media-andtelecom-predictions/2022/risc-v-open-source-cpu.html>.
- [4] 2022 Draper Prize recipients honored for contributions to the invention of RISC chips [Internet]. Washington, DC: National Academy of Engineering; 2022 Mar [cited 2022 Jul 23]. Available from: <https://www.nae.edu/276632/2022-Draper-Prize-Recipients-Honored-for-Contributions-to-the-Invention-of-RISC-Chips>.
- [5] Bailey B. A minimal RISC-V [Internet]. San Jose: Semiconductor Engineering; 2022 Jan 13 [cited 2022 Jul 13]. Available from: <https://semiengineering.com/a-minimal-risc-v/>.
- [6] Hsu J. RISC-V star rises among chip developers worldwide [Internet]. New York City: IEEE Spectrum; 2021 Apr 7 [cited 2022 Jul 5]. Available from: <https://spectrum.ieee.org/riscv-rises-among-chip-developers-worldwide>.
- [7] Haranas M. Huawei selling x86 server business due to US sanctions: report [Internet]. Westborough: CRN; 2021 Nov 4 [cited 2022 Jul 13]. Available from: <https://www.crn.com/news/data-center/huawei-selling-x86-server-business-due-to-u-s-sanctions-report>.
- [8] China opens new research institute to develop RISC-V processor project [Internet]. Beijing: TechNode; 2022 Apr 11 [cited 2022 Jul 5]. Available from: <https://technode.com/2022/04/11/china-opens-new-research-instituteto-develop-risc-v-processor-project/>.
- [9] Mitchell R. Alibaba close to porting Android to RISC-V [Internet]. Poole: Electropages; 2022 Apr 5 [cited 2022 Jul 5]. Available from: <https://www.electropages.com/blog/2022/05/alibaba-close-porting-android-risc-v>.
- [10] Martin D. One of the first RISC-V laptops may ship in September, has an NFT hook [Internet]. London: The Register; 2022 Jul 1 [cited 2022 Jul 5]. Available from: https://www.theregister.com/2022/07/01/riscv_roma_laptop/.
- [11] Analyzing the RISC-V CPU market for SIP, SoCs, AI and design starts [Internet]. Phoenix: Semico; 2021 Nov 17 [cited 2022 Jul 5]. Available from: <https://semico.com/content/analyzing-risc-v-cpu-market-sip-socs-ai-and-design-starts>.
- [12] Freund K. Esperanto launches AI accelerator with over 1000 RISC-V cores [Internet]. Jersey City: Forbes; 2021 Aug 24 [cited 2022 Jul 5]. Available from: <https://www.forbes.com/sites/karlfreund/2021/08/24/esperanto-launches-with-1000-core-risc-v-ai-accelerator/?sh=128f2b5bbded>.
- [13] SamsungMartin D., others test drive Esperanto’s 1000-core RISC-V AI chip [Internet]. London: The Register; 2022 Apr 22 [cited 2022 Jul 5]. Available from: https://www.theregister.com/2022/04/22/samsung_esperanto_riscv/.
- [14] Weiss T. Ceremorphic touts its HPC/AI silicon technology as it exits stealth [Internet]. San Diego: HPC Wire; 2022 Jan 25 [cited 2022 Jul 5]. Available from: <https://www.hpcwire.com/2022/01/25/ceremorphic-exits-stealth-touts-hpc-ai-silicon-technology/>.
- [15] Cherney MA. The newest autonomous driving chip from Intel’s Mobileye doesn’t have any of Intel’s core tech [Internet]. San Francisco: Protocol; 2022 Jan 4 [cited 2022 Jul 5]. Available from: <https://www.protocol.com/bulletins/intel-mobileye-eyeq-autonomous-driving>.
- [16] Weiss TR. Kneron unveils its first RISC-V SoC built for autonomous, assisted driving [Internet]. San Diego: Enterprise AI; 2021 Nov 3 [cited 2022 Jul 5]. Available from: <https://www.enterpriseai.news/2021/11/03/kneron-unveils-its-first-risc-v-soc-built-for-autonomous-assisted-driving/>.
- [17] Khalili J. Will Intel abandon x86 for RISC-V for its next-gen supercomputing chips? [Internet]. New York City: TechRadar; 2022 Jun 5 [cited 2022 Jul 5]. Available from: <https://www.techradar.com/news/will-intel-abandon-x86-for-risc-v-for-its-next-gen-supercomputing-chips>.
- [18] Shilov A. Researchers benchmark experimental RISC-V supercomputer [Internet]. New York City: Tom’s Hardware; 2022 Jun 11 [cited 2022 Jul 5]. Available from: <https://www.tomshardware.com/news/researchersbenchmark-experimental-risc-v-supercomputer>.