

并行计算技术的现状与未来

张明昆 2211585

2024 年 3 月 13 日

目录

1	引言	1
2	当前并行体系结构的概览	2
2.1	超级计算机	2
2.2	CPU 技术	3
2.3	GPU 技术	4
3	并行体系结构的进步与发展趋势	4
3.1	技术进步	4
3.2	发展趋势	5
4	中国与世界的超级计算机发展对比	6
4.1	中国的发展历史	6
4.2	国际发展历史	6
4.3	重要性分析	6
5	结论	6

1 引言

并行计算是一种计算模式，它利用多个处理单元同时执行计算任务，显著提高了计算速度和效率。这一概念的起源可以追溯到 20 世纪 60 年代早期，当时的科学家们开始探索使用多个处理器来解决大规模科学计算问题。自那以来，随着微电子技术的飞速发展，各种并行处理器和并行计算机系统

相继诞生，极大地推动了计算机科学、物理、化学、生物学等多个学科的发展。并行计算的重要性体现在其对解决大规模、复杂问题的能力上。在天气预报、气候变化模拟、生物信息学、材料科学、大数据分析等领域，传统的串行计算已经无法满足计算需求。并行计算技术的应用，不仅大大缩短了计算时间，还使得一些以往认为不可能解决的问题变得可行。

调研的目的在于：一方面，通过分析最新的并行计算技术，了解当前并行体系结构的技术水平和发展方向；另一方面，通过对比不同国家和地区在并行计算领域的发展历程，探讨并行计算技术对国家科技进步和经济发展的作用。此外，本调研还旨在通过分析不同并行计算架构的特点，预测未来并行计算技术的发展趋势，为相关领域的研究人员和决策者提供参考。

2 当前并行体系结构的概览

无论是在超级计算机的巨大规模还是个人计算设备的微观层面，并行计算技术都在不断推进计算能力的边界。接下来，我们将深入到这些技术在各个领域的具体应用和表现，详细分析不同并行架构在解决现实世界问题中的角色和影响，从而更全面地理解并行体系结构的进步与发展趋势。

2.1 超级计算机

Frontier: 位于美国橡树岭国家实验室的 Frontier 超级计算机，是世界上首台达到 exaflop（每秒执行一百亿亿次浮点运算）水平的系统。[1] [2] 它结合了高性能的 CPU 和 GPU，为科学研究和模拟提供前所未有的计算能力。应用场景包括气候科学、生物医药、核能模拟等领域。Frontier 利用了 AMD 的 Epyc CPU 和 Radeon Instinct GPU 的组合，整套系统包括 9,472 个 CPU 和 37,888 个 GPU，总计 CPU 内核数达 606,208 个，GPU 内核数达 8,335,360 个。[3] 展示了 CPU 和 GPU 协同工作的并行计算模型。通过这种混合架构，Frontier 实现了高效的数据处理和能源利用率，体现了异构并行计算技术的先进性。

Aurora: 位于阿贡国家实验室的 Aurora 超级计算机，自 2023 年以来，它一直是全球第二快的超级计算机。预计其性能优化后将超过 2 ExaFLOPS，成为有史以来最快的计算机。2023 年 11 月，Aurora 在 Top500 超级计算机排行榜上排名第二，性能为 585 petaFLOPS。正如 Top500 在 2023 年 11 月时事通讯中所述，Aurora 的 Top500 条目仅包含一半的硬件。一旦整个系统

上线并进行优化，Aurora 仍有望超过 2 exaFLOPS 的性能，超过 Frontier，成为 Top500 中排名第一的超级计算机。[4] [5] [6]

富岳：位于日本理化学研究所的富岳超级计算机，是目前世界上性能最强的超级计算机之一。它采用了基于 ARM 架构的自定义 CPU，强调能效比和通用计算能力，广泛应用于药物发现、灾害模拟和气候研究等领域。富岳采用了 ARM 架构的 A64FX 处理器，特别设计了高效的向量处理单元和高速的互连网络。这使得富岳在进行大规模并行处理时能够保持高效能效比，展现了面向能效的并行计算设计思想。

神威·太湖之光：位于中国无锡的神威·太湖之光，是世界上首个使用中国自主研发的处理器器的超级计算机。其主要特点是高能效比，适用于大规模数值模拟、天气预报、生命科学等多个领域。神威·太湖之光使用的是国产的申威 26010 处理器，该处理器内集成了数以千计的处理核心，采用了基于片上网络的设计，使得处理单元之间可以高效地通信，体现了大规模集成并行处理单元的设计理念。

2.2 CPU 技术

Intel 酷睿 14/13 系列：代表了 Intel 在桌面和移动处理器市场的最新进展，采用了先进的制程技术，显著提高了计算性能和能效比。酷睿 14/13 系列处理器采用了细致的能源管理机制和更高的时钟频率，以适应个人电脑、高端游戏和轻薄笔记本的需求。这些处理器通过优化的多核心和超线程技术支持，能够更高效地并行处理多任务，提升了用户的多任务处理能力和应用响应速度。

Intel Sapphire Rapids/Emerald Rapids 系列：这些处理器是针对高性能计算和数据中心市场推出的，引入了多核心设计和支持多线程技术，可以在单个芯片上并行执行多个计算任务。此外，这些处理器集成了先进的内存技术和高速 I/O 接口，显著提升了数据处理能力和效率，特别适用于云计算、大数据分析和人工智能应用。

AMD Zen 4 架构：通过采用新的制程技术和架构优化，AMD 的 Zen 4 架构在性能和能效比上取得了重大突破。Zen 4 处理器通过提升每个核心的 IPC 和增加核心数量，进一步加强了 CPU 的并行处理能力。此外，支持的高速缓存一致性和低延迟的内存访问技术，优化了核心间的数据并行传输，增强了在服务器、桌面和游戏领域的竞争力。

苹果 M3 系列处理器：在 M1 和 M2 系列的基础上，M3 系列处理器

为苹果设备带来了更高的性能和能效。采用先进的 SoC 设计，集成了多核 CPU、GPU 和 AI 加速器等处理单元，这些单元可以高效协同工作，执行复杂的并行计算任务。特别在图形处理和机器学习等领域，M3 系列处理器的并行计算能力使苹果产品在性能和能效比上保持领先。

2.3 GPU 技术

英伟达 Ada Lovelace 架构：英伟达的 Ada Lovelace 架构标志着 GPU 技术的一大进步，特别是在图形渲染和并行计算性能方面。它通过采用新的架构设计和增加计算核心，显著提高了对游戏、专业可视化应用和数据中心中 AI 计算的支持能力。Ada Lovelace 架构能够同时处理成千上万的并行线程，使其在执行深度学习和科学计算等需要高度并行处理的任務时表现出色。

AMD RDNA3 架构：AMD 的 RDNA3 架构在提升 GPU 能效比和性能方面取得了显著成就，尤其是在处理高分辨率游戏和内容创作方面。RDNA3 架构通过增加计算单元和优化内部架构，提高了并行处理图形和计算任务的能力。这一进步不仅增强了游戏体验，也为专业图形和视频处理提供了强大的支持。

Intel Ponte Vecchio：作为 Intel 针对数据中心和高性能计算市场推出的首款 GPU，Ponte Vecchio 采用了多项创新技术，包括高带宽内存和高密度计算单元。这些技术的融合，为 AI 训练和高性能计算领域提供了前所未有的计算能力和数据吞吐率。Ponte Vecchio 的设计重点是优化并行处理性能，使其在处理数据密集型任务时能够展现出卓越的能力。

3 并行体系结构的进步与发展趋势

3.1 技术进步

并行体系结构的进步主要体现在能效比、计算性能和集成度等方面。通过对比不同代的 CPU 和 GPU 架构，我们可以明显看到这些进步。

能效比：随着制程技术的进步，例如从 14nm 到 7nm 再到 5nm 甚至更先进的制程，CPU 和 GPU 的能效比得到了显著提高。这意味着在消耗相同电量的情况下，新一代的处理器能完成更多的计算任务。例如，AMD 的 Zen 3 与 Zen 4 架构相比，后者在相同功耗下提供更高的性能。

计算性能：新一代的 CPU 和 GPU 通过增加核心数、提高时钟速度和优化架构设计，实现了计算性能的大幅提升。例如，苹果 M3 系列相比 M1 系列，在保持能效的同时，性能提升显著，这得益于其更多的处理核心和更高效的架构设计。

集成度：系统级芯片（SoC）的出现，将 CPU、GPU、AI 加速器等多种处理单元集成在一个芯片上，显著提高了集成度。这不仅减少了芯片间的数据传输延迟，也提高了整体系统的性能和能效。例如，苹果的 M 系列处理器就是一个典型的 SoC 设计，其集成了多种功能，为不同的应用场景提供了强大的并行计算能力。

3.2 发展趋势

更高的能效：随着技术的发展，未来的并行体系结构将继续追求更高的能效比。这意味着新一代的处理器将在消耗更少能源的同时，提供 stronger 的计算性能。

AI 与传统计算的融合：AI 的广泛应用要求处理器具备强大的 AI 计算能力。因此，未来的并行体系结构将更加注重 AI 计算与传统计算的融合，通过专门的 AI 加速器或优化的指令集来提升 AI 应用的性能和效率。更高的集成度和系统级优化：随着技术的发展，未来的并行体系结构将实现更高的集成度，通过将 CPU、GPU、AI 加速器等多种处理单元集成在单一芯片或紧密集成的模块中，来进一步提高系统性能和能效。同时，系统级优化，如内存访问优化、数据传输优化等，将成为提升性能的关键因素。

异构并行计算模型的普及：为了适应不同计算任务的需求，异构并行计算将成为一种趋势。这种模型结合了 CPU、GPU 以及其他类型的处理器的优势，能够针对特定任务提供最优的计算资源，从而提高计算效率和性能。

量子计算与并行计算的结合：随着量子计算技术的进步，未来可能会看到量子计算与传统并行计算的结合。量子处理器在特定类型的计算任务上展现出超越传统处理器的性能，其与传统并行计算技术的结合将开启并行计算的新篇章。

并行体系结构的不断进步和发展预示着计算能力的新高度，为解决日益复杂的科学、工程和社会问题提供了强大的支撑。随着技术的发展，我们将见证更加高效、智能和集成的并行计算平台的诞生。

4 中国与世界的超级计算机发展对比

4.1 中国的发展历史

中国的超级计算机发展始于 20 世纪 90 年代末，自 2002 年银河一号问世以来，中国在超级计算机领域取得了显著进步。接下来的银河二号、天河一号、天河二号、神威·太湖之光和最新的曙光超级计算机等，标志着中国在这一领域的技术水平和研发能力的飞速发展。特别是 2016 年，神威·太湖之光成为世界上最快的超级计算机，这是中国首次在全球超级计算机排名中获得顶尖位置，展示了中国自主研发处理器的强大实力。

4.2 国际发展历史

国际上，超级计算机的发展可以追溯到上世纪 60 年代，美国 and 欧洲一直是这一领域的领导者。美国的 Cray-1（1976 年）、日本的富岳（2020 年）、美国的 Frontier（2021 年）等，都曾是世界最快的超级计算机。这些计算机在科学研究、军事和工业领域发挥了重要作用，推动了相关技术和应用的发展。

4.3 重要性分析

超级计算机在国家发展中的作用不容小觑。首先，在科学研究领域，超级计算机被用来模拟气候变化、宇宙演化、新材料的设计等复杂的自然现象，这对于提高人类对自然世界的理解具有重要意义。其次，在国防领域，超级计算机被用于核模拟、密码破解和高级战略模拟等，对于国家安全至关重要。最后，在经济方面，超级计算机在石油勘探、大数据分析和金融模拟等方面的应用，极大地提高了经济活动的效率和精确度。

5 结论

结论

参考文献

- [1] Jack Wells, "Powering the Road to National HPC Leadership," *Open-POWER Summit*, March 19, 2018, archived from the original on August 4, 2020, retrieved March 25, 2018.
- [2] "DOE Under Secretary for Science Dabbar's Exascale Update," *insideHPC*, October 9, 2020, archived from the original on October 28, 2020.
- [3] Charles Q. Choi, "The Beating Heart of the World's First Exascale Supercomputer," *IEEE Spectrum*, June 24, 2022, archived from the original on August 14, 2022, retrieved August 14, 2022.
- [4] "Intel Data Center GPU Max Series Overview," Intel, retrieved November 14, 2023.
- [5] Anton Shilov, "The Aurora Supercomputer Is Installed: 2 ExaFLOPS, Tens of Thousands of CPUs and GPUs," *www.anandtech.com*, retrieved November 14, 2023.
- [6] "Aurora - HPE Cray EX - Intel Exascale Compute Blade, Xeon CPU Max 9470 52C 2.4GHz, Intel Data Center GPU Max, Slingshot-11," *TOP500*, retrieved November 14, 2023.