

并行体系结构调研

2025/3/13

2313911 黄尚扬

目录

一. 并行体系结构的历史	3
(一) 并行体系的发展历史概述	3
二. 并行体系结构的现状	4
(一) 现行并行机体系结构分类	5
(二) 我国超算现时的创新	5
(三) 国内外超算横向对比	6
三. 并行计算的未来	6
四. 结语	7
五. 参考文献	7

一. 并行体系结构的历史

（一）并行体系的发展历史概述

1. 并行计算的发展历程

并行计算的历史可以追溯到 1960 年代，当时的大型计算机采用了多个处理单元的设计。随着计算机技术的发展，并行计算在 1970 年代和 1980 年代得到了广泛应用，尤其是在科学计算和工程计算领域。1990 年代以来，并行计算技术的发展加速，并且越来越多的商业应用开始使用并行计算。[1]

其中，并行计算机从 20 世纪 70 年代开始快速发展,到 20 世纪 80 年代出现了蓬勃发展和百家争鸣的局面，20 世纪 90 年代体系结构框架趋于统一；近 5 年来机群技术逐渐成为一个新的快速发展热点。目前，并行计算机技术日趋成熟。本节以时间为线索，简介并行计算机发展的推动力和各个发展阶段，以及各类并行机的典型代表和它们的主要特征。[2]

2. 我国超算的发展历程

在我国，超算作为并行的重要体现部分，也在时代的大潮中得到了长足发展。

1983 年，我国第一台亿次超级计算机“银河一号”研发成功，这是我国超级计算机研制的一个里程碑，也让我国成为继美国、日本后世界上第三个能够独立设计和制造超级计算机的国家。

而到了 2009 年，“天河一号”超级计算机研制成功，使得我国一跃而成为继美国之后第二个拥有千万亿次超级计算机的国家。此后，我国的超算发展迎来了一轮井喷——2010 年，我国自主研发的“星云”取得了当年的超级计算机排行榜第二名的佳绩；而到了次年，“天河-1A”计算机的运算速度成为了世界之首；而与此同时“神威·蓝光”计算机率先实现了 CPU 的国产化。可以说，中国的超算发展，已经走在了时代前列。

不过，凡事不可能一帆风顺。在我国研制超算的历程中，掩藏在熠熠生辉光芒下的，是数十年来无数科研人员的汗水，也隐含了技术上的飞速迭代。接下来笔者会就一些有代表性的超级计算机进行阐述。

前面提到的“银河”系列超级计算机，标志着我国巨型机研制能力的巨大飞跃。然而“银河”超级计算机的研制是一项极具挑战性的任务，特别是在我国计算机技术尚处于发展阶段的背景下。20 世纪 70 年代，我国气象、航空航天和石油勘探等领域对高性能计算机的需求日益迫切。例如，气象部门需要巨型计算机进行中长期天气预报，航空航天部门希望借助超级计算机降低风洞实验成本，而石油勘探依赖高性能计算机处理三维地震数据。此外，大规模数值模拟计算在工程和科学研究中的应用也逐渐增多，对计算能力的需求日益增长。然而，当时我国高性能计算机主要依赖进口，部分部门甚至需要租用外国计算机，且使用受限，这促使我国必须自主研发高性能计算机。

在研究银河-I 关键核心技术之一的**并行体系结构**时，当时主要有两个创新的方案，一个方案是双向量阵列部件单处理机系统，另一个是共享主存的双中央处理机系统。调研结果显示，该领域 75% 以上的计算任务属于向量运算，因此，采用**双向量阵列**架构能够有效提升计算效率，使其更适用于大型科学计算任务。同时，系统在设计上广泛应用了多功能计算部件、全流水化处理、并行计算及分布式结构等先进技术，以进一步提高计算性能和系统吞吐能力，同时解决元器件水平受限所带来的难题。[3]

“曙光”超级计算机是我国另一系列的超级计算机系统，始发于 90 年代。它采用了大规模并行计算架构，支持高密度计算节点，并结合高性能互连技术，显著提高了计算效率。在 93 年开始的曙光 1000 大规模并行机研制中，国家智能机中心在国内率先突破了“**驻洞路由**”这一关键技术，研制成功将大量处理机连接起来的路由芯片，为我国研制可扩展的大规模并行机探索了一条可行的道路。[4]同时，曙光系列超级计算机的 GPU 微架构采用异构计算架构，结合 CPU+GPU 协同计算，协调提升高性能计算和人工智能任务的效率。其设计特点包括大规模并行计算（SIMT 架构、多流处理器）、高带宽存储（HBM/GDDR、片上缓存优化）、高速互连（NVLink、PCIe、CXL）以及能效优化（动态频率调整、先进制程）。同时，曙光逐步采用国产 GPU，以降低对外部技术的依赖，并推动国内高性能计算的发展。

二. 并行体系结构的现状

（一）现行并行机体系结构分类

从体系结构的角度，当前并行机的体系结构可分为如下三类：

首先是机群。机群是利用标准的网络将各种普通的服务器连接起来，通过特定的方法，向用户提供更高的系统计算性能、存储性能和管理性能，同时为用户提供单一系统映象功能的计算机系统。

机群系统的良好性能可扩展性是对机群的研究灵感起源。最初，提供计算机性能的主要手段是总线带宽与 CPU 主频的提高。但这一手段对系统性能的提供十分有限。紧随其后，人们试图通过增加 CPU 个数和内存容量来提高性能；据此，出现了向量机与对称多处理机等结构。但当 CPU 的个数过多，SMP 这类型的多处理机系统可扩展性受制——而其主要瓶颈在于 CPU 访问内存的带宽并不能随着 CPU 个数的增加而有效增长。相反于 SMP，伴随 CPU 个数的增加，机群系统整体的性能几乎是线性提升的。

其次便是大规模并行处理系统。大规模并行处理系统的整个系统由多个计算结点组成，每个结点通常包含约 10 个共享存储资源的处理器。结点之间通过高性能专用网络互连，并采用分布式存储架构。系统运行专门设计的操作系统、编译系统及作业管理系统，以高效支持并行计算任务。

传统意义上的大规模并行处理系统以其专用互连网络为显著特征。与其他并行计算体系结构不同，MPP 系统的处理器或结点间互连网络通常根据特定应用需求量身定制，具有某种程度的专用并行机特性。目前，此类系统多由政府直接支持，处理器规模可扩展至数十万级别。

最后，星群系统也是现行并行体系结构的一个重要类别。它们的系统由结点构成，每个结点是一台使用共享存储或者分布共享存储技术的，包含数十、数百、乃至上千个微处理器的并行机子系统；与此同时，它采用商用机群交换机连接结点，结点间实现了分布存储，使之得以在各个结点上运行专用的结点操作系统、编译系统和作业管理系统。

（二）我国超算现时的创新

我国超算当前的创新主要体现在**芯片自主化**、**异构计算**等方面。

近年来，美国对中国在半导体技术领域的限制愈发严格，特别是自 2020 年起，EUV 光刻机的出口被严格管控。这一举措旨在将中国的逻辑芯片工艺限制在

14 纳米，DRAM 内存限制在 18 纳米，NAND 闪存限制在 128 层。[5] 现时，从晶圆制造到前道工序，再到后道工序，整个芯片生产流程中的各个环节都在逐步推进国产替代。晶圆制造方面，中国已经能够制造 300 毫米的晶圆，这对于生产 3 纳米、2 纳米的芯片而言，并不存在技术障碍。

异构计算是一种特殊的并行分布式计算系统。与同构计算（即一般说的多核理念）不同，异构计算技术使得不同类型的计算核心可以高效协同，将不同厂家、不同架构的芯片放在一个统一的计算机系统中结合 CPU+GPU+FPGA+AI 加速芯片的多架构计算模式，通过软件的调度来实现高速的 AI 计算。这种计算方式有效突破了同构计算所带来的技术瓶颈，为我国自研超算的创新作出了显著的贡献。

（三）国内外超算横向对比

在计算性能方面，全球超算 TOP500 榜单上的领先系统由中美日主导。美国的“Frontier”基于 HPE Cray EX 架构，结合了第三代 AMD EPYC CPU（针对 HPC 和 AI 优化）、AMD Instinct 250X 加速器以及 Slingshot-11 互连，在 E 级计算上率先取得突破；而我国的超算在高性能领域的总数上保持领先。[6]

其次便是架构设计方面。我国采用申威、飞腾、海光等国产 CPU，结合国产 GPU 和 AI 加速芯片，发展 CPU-GPU-FPGA 异构计算架构，以降低对国外技术的依赖；而美国作为老牌计算机强国，主要依赖 Intel、NVIDIA、AMD 等厂商的芯片，其 GPU 计算架构在 AI 计算领域保有强大优势；此外，日本的“富岳”超算采用了纯 ARM 架构，在高效低能耗方面具有强大创新性。

在互联存储层面，中美确实存在差距。美国的高效互联技术包括 NVLink、Infinity Fabric 等，采用内存池化等技术，提高多 GPU 之间的数据传输速率；而我国在这一领域逐步研发了光互连和高速互联总线技术，逐渐减小了和美国的差距；在日本和欧盟等地，更加强调低延迟的分布式存储系统，各有优势。

三. 并行计算的未来

并行计算的未来将围绕更高效的硬件架构、智能化的软件优化、异构计算的发展以及新兴计算范式展开，推动计算能力的持续提升。

并行计算现在已经处在一个全新的时代。2017 年获得图灵奖的计算机体系结构大师 Hennessey 和 Patterson 指出，无论是在指令级处理的技术还是多核处理技术，通用处理器固有的低效率和缩放比定律、摩尔定律的终结，使得处理

器架构师和设计者已经不太可能在通用的处理器上再保持显著的性能改进。[7]

一台通用性较强、性能较高的并行机，不仅要实现低延迟高带宽的互连网络接口，而且应具有可扩展性、单一系统形象和友好的并行编程环境。并行计算机水平高低也主要反映在这几项关键技术。[8]

可扩展性是高性能计算的重要设计原则和重要发展方向，涉及资源、技术升级、应用问题等多个方面。一个可扩展系统应具备计算能力、存储容量、I/O 性能同步提升的特性，并通过数据预取、多线程机制等技术隐藏延迟，优化并行效率。可扩展并行机的发展趋势是从器件级集成向微机、工作站主板甚至整机集成演进，实现与计算平台的同步升级。

此外，可扩展系统应提供单一系统形象，确保用户体验一致；而其最终目标是构建一个无性能瓶颈、平台无关、算法独立的高效并行计算系统。

四. 结语

并行计算已经经历了半个多世纪的发展，从最初的 SIMD 架构到如今的异构计算和大规模并行处理（MPP）系统，其技术不断演进，应用领域日益广泛。当前，并行体系结构的发展正朝着更高效的硬件设计、智能化的软件优化以及可扩展的计算架构迈进。随着摩尔定律的放缓，计算机体系结构正从单纯的处理器性能提升转向深度并行化、异构计算和分布式计算等新技术方向。未来，并行计算将在高性能计算、人工智能、大数据处理等领域发挥更加重要的作用，同时，如何提升并行计算的可编程性、能效比及可扩展性仍是研究的重要课题。

我国在超算领域逐步普及自研技术，缩小与美国的技术差距、甚至在某些方面超越而成为行业龙头，成果颇丰。面对不断增长的计算需求，并行计算将继续成为推动计算机技术进步的重要动力，为科学研究、工业应用及社会发展提供更加强大的算力支撑；也相信我国在不断创新之下，能够不断前进，获取超算这一代表未来领域上的工业明珠。

五. 参考文献

[1] 知乎 . 计算的原理和计算技术简史：并行计算的发展 [DB/OL].(2023-12-11)[2025-3-13].<https://zhuanlan.zhihu.com/p/671402309>.

[2] 豆丁网 . 并行计算机的发展历史 [DB/OL].(2015-5-17)[2025-3-13].<https://www.docin.com/p-1151976017.html>.

-
- [3] 记 忆 40 年：研 制 第 一 台 巨 型 计 算 机 “ 银 河 -I”[DB/OL].(2023-12-22)[2025-3-14].<http://ccfcm.org.cn/wyjsjdgs/2023-12-22/810822.shtml>.
- [4] 中国科学院计算研究所.曙光之路----曙光计算机研制与产业化十年总结 [DB/OL].(2009-9-15)[2025-3-14].https://www.ict.ac.cn/liguojiewenxuan_162523/wzlj/lgjsjg/201912/t20191227_5476624.html.
- [5] 国产芯片自主化新进展：除光刻机外，14nm 工艺已实现全面国产 [DB/OL].(2024-11-18)[2025-3-14].<https://cloud.tencent.com/developer/news/1930384>.
- [6] 世 界 最 快 的 超 级 计 算 机：前 沿 (Frontier)[DB/OL].(2022-5-30)[2025-3-14].<https://www.chaosuanwiki.com/lingxianchaosuanpinpai/shi-jie-zui-kuai-de-chao-ji-ji-suan-ji-qian-yan-Frontier.html>.
- [7] 维 基 百 科 . 支 持 向 量 机 [DB/OL].(2025-2-3)[2025-3-13].<https://zh.wikipedia.org/wiki/%E6%94%AF%E6%8C%81%E5%90%91%E9%87%8F%E6%9C%BA>.
- [8] 知 乎 . 李 国 杰 院 士：未 来 几 十 年 是 并 行 计 算 的 黄 金 时 代 [DB/OL](2020-3-26)[2025-3-13].<https://zhuanlan.zhihu.com/p/117378573>.
- [9] 中国科学院计算研究所.并行计算机的发展趋势与应用前景 [DB/OL](2009-9-11)[2025-3-13].https://www.ict.ac.cn/liguojiewenxuan_162523/wzlj/lgjjs/201912/t20191227_5476630.html.