



南開大學
Nankai University

计算机学院
并行程序设计实验报告

并行体系结构调研

姓名：蔡沅宏

学号：2213897

专业：计算机科学与技术

2024 年 3 月 17 日

目录

1 介绍	2
1.1 进行并行架构调研的意义	2
2 超算	2
2.1 Frontier	2
2.1.1 Frontier 简介	2
2.1.2 Frontier 技术细节及并行架构调查	2
3 CPU	4
3.1 Intel 酷睿 14 代	4
3.1.1 Intel 酷睿 14 代优势	4
3.1.2 Intel 酷睿 14 代并行架构的设计	4
4 中国超算发展历史	5
4.1 中国超算发展进程	5
4.2 中国在发展超算过程中遇到的困难	6
4.2.1 自主创新的难题	6
4.2.2 能耗问题严重	6
4.3 中外超算实力对比以及受到的西方势力的打压	7

1 介绍

1.1 进行并行架构调研的意义

并行体系结构的优化对计算机领域具有深远的意义，这不仅在学术界有广泛的研究和讨论，也在实际应用中展现出巨大的潜力和价值。根据调研结果总结，并行体系结构的意义可以大致归纳为以下几点：

性能提升：并行体系结构的进步可以显著提高计算机系统的性能。通过并行处理多个任务或数据，可以将计算任务分解为更小的子任务并同时处理，从而加速计算过程。例如，研究表明，采用并行计算技术可以在处理大规模数据集或复杂计算任务时获得数倍乃至数十倍的性能提升 [2]。

扩展性和可伸缩性：随着数据规模和计算需求的不断增长，传统的串行计算模式已经无法满足需求。并行体系结构的进步可以实现系统的扩展性和可伸缩性，使其能够处理更大规模的数据和更复杂的计算任务。研究表明，采用并行计算技术可以轻松应对大规模并行任务，不会因为系统规模的增加而导致性能下降 [1]。

能效提升：并行体系结构的进步还可以提高系统的能效。通过有效利用系统资源并并行处理多个任务，可以降低系统的能耗，并在单位能量消耗下实现更多的计算量。研究表明，并行计算技术在能源密集型应用中具有显著的能效优势，可以降低能源消耗并延长系统的工作时间 [3]。

应用领域拓展：并行体系结构的进步也为更广泛的应用领域带来了可能。除了传统的科学计算和工程应用外，现代社会中的各个领域都在不断涌现出对大规模数据处理和复杂计算的需求，如人工智能、生物信息学、物联网等。并行计算技术为这些应用提供了强大的支持，并推动了这些领域的发展和创新。

综上所述，并行体系结构的优化，可以实现计算系统性能的显著提升、系统的扩展性和可伸缩性、能效的提升以及应用领域的拓展。这些都为计算机领域的发展和應用带来了重要的意义和价值。因此，进行并行架构的调研，了解世界最新的并行结构是有必要的。

本篇调研将从世界最新超算和 CPU 入手，探索并行架构研究的最新结构，并分析其对计算机系统性能的提升程度。同时，本篇文章还将调研中国超算发展历史，通过对比中外超算发展历程，用图表与数据，全面展现如今超算领域的激烈竞争，体现国之重器——超算的重要性。

2 超算

2.1 Frontier

2.1.1 Frontier 简介

Frontier, 美国橡树岭国家实验室超级计算机。内部使用的是 AMD 在 2021 年发布的第三代 EPYC (霄龙) 服务器处理器。采用了最新的技术，Frontier 能耗同样也是排行榜中最优秀的，每瓦浮点运算能力达到 522.3 亿亿次。

在近几年的 Top500 榜单中，Frontier 已连续多次成为榜单第一名。如图2.1所示。因此，调研 Frontier 的相关技术和并行架构是有意义的。

2.1.2 Frontier 技术细节及并行架构调查

首先，Frontier 采用了超大规模集群系统，并采用了多种不同的计算处理器结构，包括 AMD 的 EPYC 处理器和 NVIDIA 的 A100 Tensor Core GPU。这种多结构的设计可以充分利用不同处理器的

Rank	System	Cores	Rmax (PFlop/s)	Rpeak (PFlop/s)	Power (kW)
1	Frontier - HPE Cray EX255a, AMD Optimized 3rd Generation EPYC 64C 20Hz, AMD Instinct M250X, Slingshot-11, HPE DOE/SC/Dak Ridge National Laboratory United States	8,730,112	1,102.00	1,685.65	21,100
2	Supercomputer Fugaku - Supercomputer Fugaku, A64FX 48C 2.20Hz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442.01	537.21	29,899
3	LUMI - HPE Cray EX255a, AMD Optimized 3rd Generation EPYC 64C 20Hz, AMD Instinct M250X, Slingshot-11, HPE EuroHPC/CSC Finland	1,110,144	151.90	214.35	2,942
4	Summit - IBM Power System AC922, IBM POWER9 Z1C 3.070Hz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM DOE/SC/Dak Ridge National Laboratory United States	2,414,592	148.60	200.79	10,096
5	Sierra - IBM Power System AC922, IBM POWER9 Z1C 3.10Hz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM / NVIDIA / Mellanox DOE/NNSA/LNL United States	1,572,480	94.64	125.71	7,438

(a) Top500 2021.6 榜单

Rank	System	Cores	Rmax (PFlop/s)	Rpeak (PFlop/s)	Power (kW)
1	Frontier - HPE Cray EX255a, AMD Optimized 3rd Generation EPYC 64C 20Hz, AMD Instinct M250X, Slingshot-11, HPE DOE/SC/Dak Ridge National Laboratory United States	8,699,904	1,194.00	1,679.82	22,703
2	Supercomputer Fugaku - Supercomputer Fugaku, A64FX 48C 2.20Hz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442.01	537.21	29,899
3	LUMI - HPE Cray EX255a, AMD Optimized 3rd Generation EPYC 64C 20Hz, AMD Instinct M250X, Slingshot-11, HPE EuroHPC/CSC Finland	2,220,288	309.10	428.70	6,016
4	Leonardo - BullSequana XH2000, Xeon Platinum 8358 32C 2.40Hz, NVIDIA A100 SXM4 64 GB, Quad-rail NVIDIA HSR100 Infiniband, EVIDEN EuroHPC/CINECA Italy	1,824,768	238.70	304.47	7,404
5	Summit - IBM Power System AC922, IBM POWER9 Z1C 3.070Hz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM DOE/SC/Dak Ridge National Laboratory United States	2,414,592	148.60	200.79	10,096

(c) Top500 2022.6 榜单

Rank	System	Cores	Rmax (PFlop/s)	Rpeak (PFlop/s)	Power (kW)
1	Frontier - HPE Cray EX255a, AMD Optimized 3rd Generation EPYC 64C 20Hz, AMD Instinct M250X, Slingshot-11, HPE DOE/SC/Dak Ridge National Laboratory United States	8,730,112	1,102.00	1,685.65	21,100
2	Supercomputer Fugaku - Supercomputer Fugaku, A64FX 48C 2.20Hz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442.01	537.21	29,899
3	LUMI - HPE Cray EX255a, AMD Optimized 3rd Generation EPYC 64C 20Hz, AMD Instinct M250X, Slingshot-11, HPE EuroHPC/CSC Finland	2,220,288	309.10	428.70	6,016
4	Leonardo - BullSequana XH2000, Xeon Platinum 8358 32C 2.40Hz, NVIDIA A100 SXM4 64 GB, Quad-rail NVIDIA HSR100 Infiniband, EVIDEN EuroHPC/CINECA Italy	1,824,768	238.70	304.47	7,404
5	Summit - IBM Power System AC922, IBM POWER9 Z1C 3.070Hz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM DOE/SC/Dak Ridge National Laboratory United States	2,414,592	148.60	200.79	10,096

(b) Top500 2021.11 榜单

Rank	System	Cores	Rmax (PFlop/s)	Rpeak (PFlop/s)	Power (kW)
1	Frontier - HPE Cray EX255a, AMD Optimized 3rd Generation EPYC 64C 20Hz, AMD Instinct M250X, Slingshot-11, HPE DOE/SC/Dak Ridge National Laboratory United States	8,699,904	1,194.00	1,679.82	22,703
2	Aurora - HPE Cray EX - Intel Exascale Compute Blade, Xeon CPU Max N970 32C 2.40Hz, Intel Data Center GPU Max, Slingshot-11, Intel DOE/SC/Argonne National Laboratory United States	4,742,808	585.34	1,059.33	24,487
3	Eagle - Microsoft NDV5, Xeon Platinum 8480C 48C 20Hz, NVIDIA H100, NVIDIA Infiniband NDR, Microsoft Microsoft Azure United States	1,123,200	561.20	846.84	
4	Supercomputer Fugaku - Supercomputer Fugaku, A64FX 48C 2.20Hz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442.01	537.21	29,899
5	LUMI - HPE Cray EX255a, AMD Optimized 3rd Generation EPYC 64C 20Hz, AMD Instinct M250X, Slingshot-11, HPE EuroHPC/CSC Finland	2,220,288	309.10	428.70	6,016

(d) Top500 2022.11 榜单

图 2.1: 近两年 Frontier 在超算领域的突出表现

优势，在各种应用场景中达到更高的性能和效率。

其次，Frontier 采用了高性能网络架构，包括多种不同类型的互连网络，如 InfiniBand、Ethernet 和 Aries 网络。这种设计可以提供高带宽和低延迟的通信能力，支持大规模并行计算和数据传输。当谈论高性能计算和数据中心网络时，InfiniBand、Ethernet 和 Aries 网络是三种常见的技术。由于这样的多种网络架构的结合设计，让 Frontier 拥有了极强的并行计算能力，这三种网络架构各具特点，以下是对于三种网络架构的简要分析。

InfiniBand:

InfiniBand 是一种高性能、低延迟的网络技术，最初设计用于连接服务器、存储设备和网络设备。它提供了高速数据传输和低延迟的特性，使其成为高性能计算（HPC）和数据中心领域的首选之一。InfiniBand 支持多种拓扑结构，包括点对点、环形和网格拓扑，这使得它非常灵活，可以根据不同的需求进行配置和部署。它的主要特点包括双向数据传输、RDMA（远程直接内存访问）、低延迟和高带宽。这使得 InfiniBand 非常适合需要大量数据传输和低延迟的应用程序，例如科学计算、金融交易和大数据分析。

Ethernet:

Ethernet 是最常见的局域网（LAN）技术之一，被广泛应用于企业和家庭网络中。它是一种成熟的技术，具有广泛的硬件和软件支持，并且成本相对较低。Ethernet 最初设计用于传输数据，但随着技术的发展，现在它也支持音频、视频和其他多媒体数据的传输。传统以太网通常以 1、10、25、40、100 或更高的速度运行，速度取决于硬件和网络设备的能力。最新的 Ethernet 标准还支持 RDMA 和低延迟特性，使其在一些高性能计算场景下成为一种备选方案。

Aries 网络:

Aries 网络是由 Cray（现为 HPE）开发的一种高性能互连网络技术，用于连接超级计算机中的节点和存储资源。它被设计用于支持超大规模、高性能的计算和数据处理任务。Aries 网络采用了自适应路由和虚拟化技术，以提供低延迟、高带宽和高可靠性的通信。它还支持 RDMA，允许节点直接访问彼此的内存，从而加速数据传输和处理。与传统的以太网和 InfiniBand 相比，Aries 网络通常用于超级计算机和高性能计算环境中，其中需要极高的性能和可扩展性。总的来说，InfiniBand、Ethernet 和 Aries 网络都是用于构建高性能计算和数据中心网络的重要技术，每种技术都有其独特的特点和适用场景。选择其中一种技术通常取决于具体的应用需求、性能要求和预算限制。

第三，Frontier 采用了先进的存储系统设计，包括高速固态存储器、磁盘存储器以及高速缓存存储器。这种设计可以提供更高的存储容量和更快的数据读写速度，支持大规模数据处理和深度学习等应用。

第四，最后，Frontier 还采用了先进的软件和编程工具，包括多种不同的并行编程模型例如 MPI、OpenMP、CUDA、OpenACC（用于 C、C++ 和 Fortran 代码）和优化工具，以及高效的作业调度和管理系统。这种设计可以提高应用程序的性能和可扩展性，支持更复杂和更大规模的科学和工程计算。

3 CPU

3.1 Intel 酷睿 14 代

3.1.1 Intel 酷睿 14 代优势

首先，英特尔酷睿第 14 代处理器采用了全新的 10nm 工艺制程，相较于上一代产品，它能够集成更多的晶体管，并提供更高的频率和更低的功耗。在实际测试中，它展现出了卓越的性能，在多核处理能力和单核处理速度上都有了显著的提升。这意味着用户在进行多重任务时，处理器能够更快速地完成计算，并且在执行单个任务时也能够更加迅捷。

其次，英特尔酷睿第 14 代处理器还支持深度学习和人工智能等领域的计算任务。在这个信息爆炸的时代，海量的数据正在以前所未有的速度不断涌现。为了更好地处理和分析这些数据，计算机需要具备强大的计算能力和智能化的算法。酷睿第 14 代处理器通过优化架构和增强 AI 指令集，使得它在实现这些任务时表现出色。它可以支持更快的数据处理速度和更高的并行计算能力，为数据科学家和机器学习工程师提供更好的工具和平台。

此外，英特尔酷睿第 14 代处理器在能效方面也有了很大的突破。在过去，人们普遍认为性能和功耗之间存在着天然的矛盾。然而，英特尔公司通过全新的工艺和架构设计，使得酷睿第 14 代处理器能够实现更高的性能和更低的功耗，并在整个处理过程中保持较低的温度。这使得它在长时间高负载工作的情况下也能够始终保持良好的稳定性和效率。

3.1.2 Intel 酷睿 14 代并行架构的设计

酷睿 14 代芯片延续了自 12 代酷睿以来英特尔推出的全新混合异构设计，在此设计下处理器往往会搭载更多的核心数，且通过性能核（P 核）与能效核（E 核）协同工作的灵活调度机制，全面提升 CPU 和性能和能效表现。这样的并行设计让酷睿芯片在指令执行上有着更快的处理速度。并且，酷睿 14 代芯片的微架构依旧是 Raptor Cover Refresh，这样的微架构让 Intel 最新的几代芯片都拥有很强的并行计算的能力。

就 14 代酷睿处理器本身而言，英特尔优化了 CPU 的多线程处理能力，多线程性能提升 18%。对于游戏玩家来说，英特尔通过线程管理器和架构优化，提高了 CPU 的并行处理能力，提升了游戏对

于 CPU 依赖部分的性能，还有更低延迟的连接，或者是超频能力的提升。另外，14 代酷睿还具备一定的 AI 性能，官方甚至提供了具有智能自动超频功能的 AI Assist 超频助手应用。

Raptor Cover Refresh Raptor Cache Refresh 架构是一种高效的缓存刷新技术，旨在改善计算机系统中缓存的性能和可靠性。这种架构最初由学术界提出，后来被业界广泛应用于各种计算机系统和芯片设计中。

下面是 Raptor Cache Refresh 架构的主要特点和工作原理：

1. 增量刷新：Raptor Cache Refresh 采用增量刷新的方式更新缓存行。与传统的全局刷新相比，增量刷新只更新被修改的缓存行，而不是整个缓存，从而减少了刷新操作的成本和延迟。

2. 并行刷新：该架构允许多个缓存行同时进行刷新操作，从而利用了并行处理的优势，提高了刷新效率。

3. 动态调整刷新速率：Raptor Cache Refresh 能够根据系统的负载和性能需求动态调整刷新速率。在系统负载较轻的情况下，可以降低刷新频率以节省能量和提高性能；而在负载较重的情况下，则可以加快刷新速率以确保缓存数据的一致性。

4. 错误处理和容错：该架构还包括错误处理和容错机制，以确保在发生错误时能够及时检测并进行恢复，从而提高了系统的可靠性和稳定性。

4 中国超算发展历史

4.1 中国超算发展进程

中国在超级计算机方面发展迅速，已跃升到国际先进水平国家当中。同时，中国也是第一个以发展中国家的身份制造了超级计算机的国家。

中国在 1983 年就研制出第一台超级计算机银河一号，使中国成为继美国、日本之后第三个能独立设计和研制超级计算机的国家。中国以国产微处理器为基础制造出本国第一台超级计算机名为“神威蓝光”，在 2019 年 11 月 TOP500 组织发布的世界超级计算机 500 强榜单中，中国占据了 227 个（这是 Top500 榜单中国上榜数量最多的记录），神威·太湖之光超级计算机位居榜单第三位，天河二号超级计算机位居第四位。

中国的计算机行业起步并不算晚，通过学习苏联的计算机技术，1958 年 8 月 1 日中国第一台数字电子计算机——103 机诞生。

进入 70 年代，中国对于超级计算机的需求日益激增，中长期天气预报、模拟风洞实验、三维地震数据处理、以至于新武器的开发和航天事业都对计算能力提出了新的要求。为此中国开始了对超级计算机的研发，并于 1983 年 12 月 4 日研制成功银河一号超级计算机。这是我国高速计算机研制的一个里程碑，也让我国成为世界上第三个能够独立设计和制造超级计算机的国家。之后继续成功研发了银河二号、银河三号、银河四号为系列的银河超级计算机，使我国成为世界上少数几个能发布 5 至 7 天中期数值天气预报的国家之一。并于 1992 年研制成功曙光一号超级计算机，在发展银河和曙光系列同时，中国发现由于向量型计算机自身的缺陷很难继续发展，因此需要发展并行型计算机，于是中国开始研发神威超级计算机，并在神威超级计算机基础上研制了神威蓝光超级计算机。2002 年联想集团研发成功深腾 1800 型超级计算机，并开始发展深腾系列超级计算机。

2009 年，成功研制“天河一号”千万亿次超级计算机，使我国成为世界上第二个成功研制千万亿次超级计算机的国家。2010 年，“星云”千万亿次计算机在第三十五届超级计算机 TOP500 排行榜荣获第二名的佳绩，进入世界超级计算机的前三甲。半年后“天河-1A”直接成为排行榜第一，“神威·蓝光”率先完成 CPU 国产化。2018 年，我国有三台超级计算机进入 E 级超算研发，分别是曙光、天河和

神威，并逐步实现 CPU 和加速器的全国产化。我们刚说到的 E 级是指秒钟运算一百亿亿次，从这就可以看出超级计算机的发展之迅速。2021 年，第五十八届全球超级计算机 TOP500 排行榜中，中国超级计算机有 173 台进入榜单，占比 34.6%，数量上超过了美国。2022 年，神威太湖之光进入 TOP500 榜单的第六名。

4.2 中国在发展超算过程中遇到的困难

4.2.1 自主创新的难题

中国超级计算机的自主创新之路并不平坦。2010 年 11 月，当“天河一号”成为世界最快的超级计算机，其核心部件 90% 以上却依然采用了美国英特尔公司和 AMD 公司最先进的 CPU（中央处理器）和 GPU（图形加速器）。连续六次登顶世界超级计算机 500 强之首的“天河二号”，其硬件系统主要部件的计算阵列全部采用美国商用微处理器。2010 年曾排名世界第二的中国“曙光星云”超级计算机，核心部件全部是由国外英特尔和英伟达芯片构建的。在中国已成为世界第二大经济体的今天，国产超级计算机系统核心部件仍大量使用国外进口芯片就不免受到公众的质疑。事实上，以集成创新为主的模式最终只能仿制别人、依靠别人，引进消化吸收再创新的模式最后会落入“落后-引进-再落后-再引进”的陷阱。2015 年，美国政府禁止了英特尔等企业向中国出口超级计算机有关的设施与技术；2018 年，美国商务部又签署了禁购令，要求所有美国公司不得向中国企业出口包括芯片、软件、操作系统等商品。这对从事中国超级计算机研制的一些单位产生了不利影响，使已经运行的某型超级计算机系统未能继续升级，甚至导致某项百亿亿次级计算系统研制进度一再推迟。

如今，中国的最强超算“神威·太湖之光”使用自主研发的国产 CPU 作为核心部件，开始探索“完全原始创新”研发模式。它是世界上首台运行速度超过十亿亿次的超级计算机，也是中国第一台全部采用国产处理器构建的超级计算机。然而，“神威·太湖之光”使用的国产“申威 26010”众核处理器与当今国际最先进的芯片相比，仍然有不小差距。著名超级计算机专家、中国科学院院士周兴铭坦言，“目前国内制造的 CPU 从整体上来看性能低、功耗高，物理设计比国外差一代以上，工艺差两代以上”。

过去 40 多年，国家在超级计算机芯片技术原始创新等方面的研究和投入都远远不足，依靠引进和集成去追求“世界第一”并不是长久之计，中国超级计算机自主创新任重道远。

4.2.2 能耗问题严重

随着超级计算机的不断升级，大规模并行计算系统需要花费更多能量，其能源消耗呈逐年增长趋势。能耗也成了与计算速度同等重要的度量指标，国际业界科学家从 2007 年起发起与 TOP500 同步的 Green500 排行榜作为国际超级计算机能耗水平的权威评价。美国 2012 年 10 月和 2018 年 6 月排名世界最快超级计算机的“泰坦”“顶点”都是相对能耗较低的机器，其制冷技术也发展到了更先进的浸入式液冷模式。

“神威·太湖之光”和“天河二号”超级计算机使用大规模通风散热结合水力空调系统用于制冷，总体上不利于降低能耗。“神威·太湖之光”一年的用电量达到 15 兆瓦，相当于 3 个清华大学的用电量。“天河二号”能耗高达 24 兆瓦，是“天河一号”的 6 倍，问世后即成为当时世界超级计算机中功耗最大的机器。“天河二号”年耗电量约为 2 亿度，一年仅电费就要 1 亿元人民币，全速运算的话，电费更高达 1.5 亿。此外，为了应对全机散热系统本身因为高负荷产生的高温，广州市政府在超算中心 3 公里外专门为“天河二号”建设了一个冷水厂，每天不间断运送 8 摄氏度的水为散热系统降温，这又增加了相当多的能源消耗，一度引起国内舆论批评。民众甚至调侃，未来可以把超级计算机建到核电站旁边。从现实角度考虑，为超级计算机配套一个核电站是不可能的，但如果按现有条件发展，中

参考文献

- [1] Franck Cappello and et al. Grid' 5000: a large scale and highly reconfigurable experimental grid testbed. *International Journal of High Performance Computing Applications*, 20(4):481–494, 2009.
- [2] Ian Foster and Carl Kesselman. *The Grid: Blueprint for a New Computing Infrastructure*. Morgan Kaufmann, 1999.
- [3] et al. Shen, Y. Green supercomputing: A review of power efficiency in high-performance computing. *ACM Computing Surveys (CSUR)*, 51(1):1–36, 2018.