



南開大學
Nankai University

计算机学院
并行程序设计第一次报告

并行体系结构调研报告

姓名：宋卓伦

学号：2311095

专业：计算机科学与技术

2025 年 3 月 18 日，南开大学计算机学院，天津

Abstract

当前,海量数据需求与日俱增,研发超级计算机并借助并行体系结构提升运算速率,已然成为众多科学家及相关专业人士的关键研究焦点。我国人口基数庞大,每时每刻均会生成海量数据。构建我国自主可控的超级计算机体系,优化并行计算架构,对于推动我国科技、经济等多领域发展意义非凡,有助于提升国家在数据处理、科学研究等方面的核心竞争力,助力国家长远战略目标的达成。本调研报告旨在调研我国超算发展和调研国际上的知名超算,分析先进体系的特点,了解我们当下的工作。

关键词: 超级计算机, 并行体系架构

目录

1 超级计算机概述	2
1.1 超算的定义	2
1.2 超算的起源	2
2 超算代表技术分析	2
2.1 并行体系结构	2
2.1.1 分布式内存系统	2
2.1.2 共享内存模型系统	3
2.2 对代表性并行体系结构的分析	4
3 我国超算的发展技术	5
3.1 我国超算的发展历史	5
3.2 我国著名的超级计算机	6
4 国际上领先的超算	6
4.1 酋长岩	6
4.2 富岳	6
4.3 Frontier	7
5 国内外超算技术对比	8
5.1 重要数据一览	8
5.2 数据的分析	8
6 对我国超算技术的展望	8

1 超级计算机概述

1.1 超算的定义

超级计算机 (Supercomputer, 简称超算) [1] 是能够执行一些复杂任务的高速运算的计算机。

超级计算机是计算能力达到 PFLOPS (千万亿次) 或 EFLOPS (百亿亿次) 级别的并行计算系统, 通过任务分片 + 多处理器协同将复杂计算拆解为纳秒级同步的并行任务流。其核心突破在于用三维 torus 网络实现处理器间 0.1 微秒延迟通信, 并通过定制数据结构 (如块分解 + 流水线预取), 让每个计算单元始终处于 "数据 - 指令" 饱和状态, 算力利用率比普通服务器集群提升 400% 以上, 利用率可达 80% 以上。

1.2 超算的起源

"超级计算" (super-computing) 概念 1929 年首次出现于《纽约世界报》有关 IBM 为哥伦比亚大学造大型制表机的新闻。20 世纪 60 年代, 西摩·克雷在控制数据公司设计出领先市场的超级计算机。70 年代他创立克雷研究公司, 1985 - 1990 年凭新设计主导超级计算机市场。但是, 80 年代小型计算机兴起, 众多对手参与竞争, 90 年代中期不少因市场冲击消失。如今, 超级计算机由 IBM、惠普等大公司专门设计, 虽这些公司借并购积累经验, 但克雷研究仍是该领域巨头之一。

2 超算代表技术分析

2.1 并行体系结构

并行体系结构是超算的核心。选择适当的并行体系结构可以优化超算的计算能力。

2.1.1 分布式内存系统

这种架构将多个计算节点的内存资源通过网络连接起来形成统一的内存池。与传统的集中式内存系统不同, 它允许数据分布在多个节点的内存中, 但是与共享内存有着如无全局地址空间¹, 依赖消息传递等本质区别, 从而支持更大规模的数据处理和更高的并发性能。这种系统通常用于大数据分析等需要处理海量数据或高并发请求的场景。

该系统采用 "内存网格化" 设计理念, 通过数据分布算法实现跨节点存储。其核心流程包含:

1. **分片策略**: 基于哈希分区或范围分片, 将数据集拆解为 256KB-1GB 不等的子块, 每个分片映射至特定节点内存;
2. **跨节点寻址**: 计算节点发起的数据请求通过元数据服务定位目标分片, 利用 RDMA 协议直接远程读取对端内存, 延迟控制在 1-10 微秒;
3. **一致性保障**: 为了确保数据的一致性, 系统会定期或根据需要进行数据同步, 确保所有节点上的数据保持一致;
4. **动态负载均衡**: 通过监控各节点内存利用率自动迁移热点分片, 保障节点负载差在一定数值下。

系统架构核心组件还包括计算节点、高速互连网络、内存管理模块、资源调度、一致性协议等。

¹ 分布式内存系统通常没有物理意义上的全局地址空间, 但某些架构 (如 PGAS) 提供逻辑上的全局地址访问能力。

2.1.2 共享内存模型系统

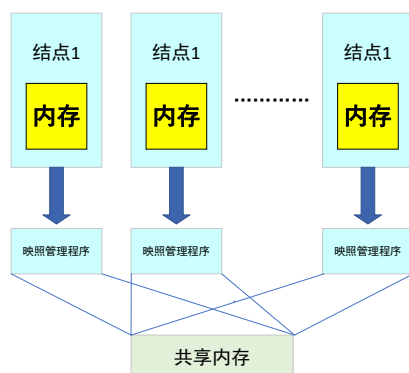


图 2.1: 共享内存分布模型

如图2.1所示，这是一个简单的共享模型的分布网络。实际情况相对复杂。多处理器环境的内存共享模型主要有 3 种：[2]

- Uniform Memory Access (UMA)
- Non-uniform Memory Access (NUMA)
- Cache-only Memory Access (COMA)

UMA （一致性内存访问，图2.2）架构中，多处理器通过共享总线访问统一内存，访问延迟均等，但存在总线争用冲突和带宽瓶颈问题，总线带宽有限导致延迟。多个处理器会争用 memory controller 造成冲突。引入对称多处理器（SMP，图2.3）变体强调处理器无主从之分，多处理器共享同一个内存，适用于 PC/移动端并行计算，但因集中式内存限制，难以扩展至超大规模服务器场景（如云计算）。

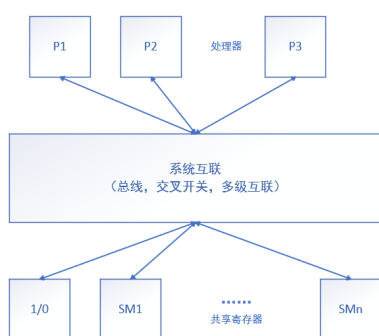


图 2.2: UMA

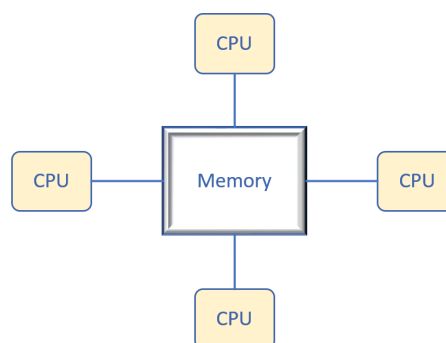


图 2.3: SMP

NUMA （非一致性内存访问，图2.4）每个处理器都有一个本地内存，且处理器可以访问其他处理器的本地内存（远端内存）。也可以一小组 CPU 一起访问它们自己的本地内存。存在多组 CPU 和它们的内存组时，每组 CPU 和内存组就构成一个 NUMA 节点（node）。NUMA 的不足（图2.5）是处

理器访问远端内存时会产生延迟。CPU 访问本地内存时速度很快，当本地内存不够用时可以使用远端内存。然而访问远端内存时延开销比较高。

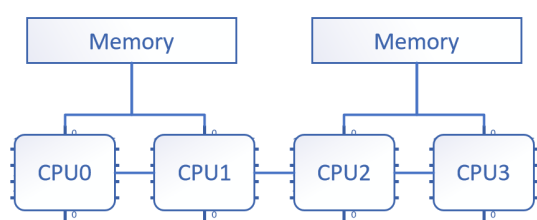


图 2.4: NUMA

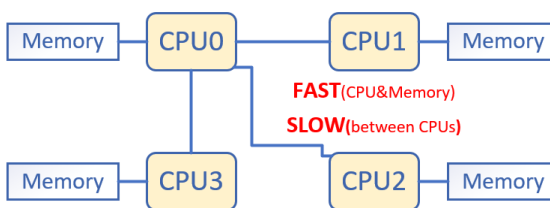


图 2.5: NUMA 的缺点

COMA 与 ccNUMA (Cache Coherent NUMA, 缓存一致性 NUMA) 2.6 的缓存一致性协议不同。COMA 通过全局缓存目录管理内存，可以看成是专用的 NUMA，将 NUMA 中的分布式内存用高速缓存来取代。全局地址空间由高速缓存组成。访问远端的高速缓存借助分布式高速缓存目录进行。

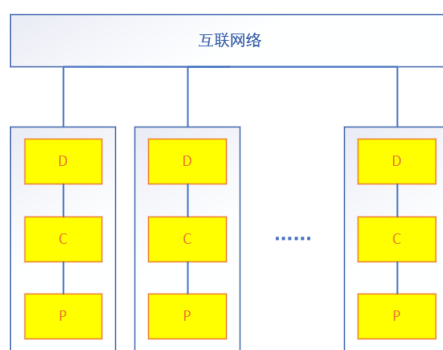


图 2.6: COMA，其中 P: 处理机; C: 高速缓存 (cache); D: 目录

2.2 对代表性并行体系结构的分析

除了上面初步提到的 SMP，我们再陈述两个代表性的体系结构。

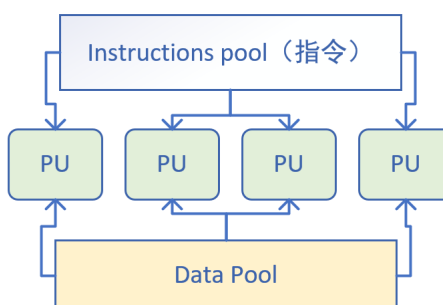


图 2.7: SIMD 单指令多数据

GPU 并行架构 GPU 并行架构的特点基于众核设计（如 NVIDIA 的 CUDA 架构），支持大规模数据级并行（DLP）。通过数千个轻量级线程同时处理数据，适合图形渲染、深度学习训练。

GPU 并行架构最重要的特性在于单指令多数据 (Single Instruction Multiple Data, 图2.7), 在单周期内对多数据进行相同计算操作。这使得 GPU 下的并行架构具有高能效比, 同时计算密度较传统 CPU 提升 1-2 个数量级。目前在现代处理器应用:

1. Intel: Intel 处理器使用的 SIMD 技术包括 MMX、SSE (SSE、SSE2、SSE3、SSSE3、SSE4)、AVX (AVX、AVX2、AVX-512) 指令集。这些指令集涵盖了多种数据类型的并行操作, 包括**整数、小数及双精度浮点数**。
2. AMD: 除了部分沿用 Intel 的指令集外, AMD 也开发了自己的**扩展指令集**如 3DNow。
3. ARM: 在移动和嵌入式设备领域, **ARM 的 NEON 技术**也是一种成熟的 SIMD 实现方式, 广泛应用于各种高性能低功耗设备中。

尽管 AMD/NVIDIA GPU 在通用计算领域占据主导, 但我国超算若长期依赖进口加速器将面临供应链风险。神威众核处理器的成功证明异构架构自主化可行性, 建议未来结合存算一体技术 (如忆阻器) 突破“内存墙”限制, 构建差异化竞争优势。

云计算 云计算是以分布式计算为核心, 通过虚拟化技术整合网络资源 (存储、算力、网络), 按需向用户提供弹性服务的体系。其发展历经三个阶段:

1. **早期阶段**: 以网格计算为基础, 通过任务分解与多服务器并行处理 (如结果合并) 实现高效运算, 可在秒级完成海量数据处理;
2. **技术融合期**: 融合分布式计算、负载均衡、并行计算、网络存储及热备冗余等技术, 形成资源池化、服务可配置的完整架构;
3. **现代云服务**: 构建于计算机网络 (如互联网) 之上, 核心能力扩展为“资源即服务”, 通过多种部署模式, 实现个性化服务交付。

其中的关键技术支撑在于: **虚拟化** (如容器化) 与分布式存储实现资源解耦与高可用性; **自动化编排** (如弹性伸缩、Serverless) 提升资源利用率; 软件定义网络 (SDN) 与**安全共担模型**保障服务可靠性与合规性。核心价值包括**敏捷性**的分钟级资源供给, 支持业务快速迭代; 提高**经济性**, 按需付费降低 IT 成本, 资源利用率提升 30%-50%; **智能演进** AI 驱动资源调度、云边协同优化实时响应, 绿色技术降低 PUE 至 1.2 以下。

3 我国超算的发展技术

我国的超算与西方国家相比起步较晚, 但是发展迅速, 已经运用到了多个领域之中。

3.1 我国超算的发展历史

1960 年, 夏培肃团队研制新中国首台自主电子管计算机 107 机, 支撑教学科研与国防计算。

1958 年中科院 109 厂在王守武等专家主导下突破高频合金扩散晶体管技术, 1960 年完成 109 乙型机核心器件国产化。1964 年哈军工研制首台全晶体管计算机 441B-I, 其高可靠性推动气象、军工领域规模化应用。1965 年研制的 109/119 计算机承担氢弹关键运算, 被誉为“功勋机”。在集成电路领域, 1965 年首块国产集成电路仅晚于美国 5 年, 1972 年实现大规模集成电路自主突破显著缩短技术代差。

3.2 我国著名的超级计算机

由于国外技术封锁和国内技术落后的原因，中国首台超算银河-I 在研制之初，就确定了“走一条自力更生、学习与独创相结合，起点跟得上当时国际水平的新路子”。

银河一号 （1983 年）基于美国 Cray-1 架构，通过原始与集成创新融合攻克双向量阵列体系、素数模双总线存储控制等核心技术，实现了高性能与低成本平衡。其技术路径体现改革开放初期“引进-消化-吸收”策略，该模式延伸至银河系列后续机型。同期中科院与联想通过增加“再创新”，推动超算体系从大规模并行向可扩展共享存储并行跨越，研制出曙光、深腾等机型，形成技术追赶的复合创新范式。

天河一号 中国首台千兆次超算“天河一号”于 2008 年启动研制，2010 年 11 月登顶国际 TOP500 榜单，2012 年位列全球第五。该系统自 2010 年投入应用后，支撑航天工程、气象气候预报、海洋环境模拟等重大领域，并应用于雾霾监测治理，彰显中国超算技术的突破性进展。

神威·太湖之光 神威太湖之光是由国家并行计算机工程技术研究中心研制的超级计算机。安装了 40960 个中国自主研发的“神威 26010”众核处理器，该众核处理器采用 64 位自主申威指令系统，峰值性能为 12.5 亿亿次 / 秒，持续性能为 9.3 亿亿次 / 秒。神威·太湖之光超级计算机具有完全自主知识产权的“神威 26010”芯片所组装的小型服务器，一颗芯片的计算能力就能够达到 3 万亿次 / 秒。

4 国际上领先的超算

4.1 酋长岩

酋长岩 (El Capitan) 是一台百亿亿级超级计算机，位于美国加利福尼亚州利弗莫尔市的劳伦斯利弗莫尔国家实验室，该系统基于 Cray EX Shasta 架构设计，2024 年投入运行。酋长岩在 2024 年 11 月发布的第 64 届全球超级计算机 500 强榜单 [3] 中超越 Frontier，成为世界上最快的超级计算机。酋长岩是美国部署的第三个百亿亿级系统，其主要用途是支持美国国家核安全管理局的核武库管理计划。

从硬件规格来看，酋长岩搭载了超过 1,100 万个处理器和图形核心，采用 44,544 个 AMD 最新型的 MI300A 集成型处理器。这些处理器是 AMD 公司专为超级计算机打造的重要突破，结合了 AMD EPYC Genoa 中央处理器、AMD CDNA3 显示卡与计算内存于单一芯片中。

4.2 富岳



图 4.8: 日本富岳计算机，初投入时“碾压”中美

富岳 (Supercomputer Fugaku, 图4.8) 是富士通与日本理化学研究所共同开发的超级计算机，作为“京”的后继机型。2014 年开始研发，2021 年正式激活。富岳是全球首度夺冠的 ARM 架构超级计算机，采用富士通 48 核心 A64FX SoC，与过往超级计算机大多采用的 Intel 或 AMD 的 x86、x64 主流平台不同。富岳共有 158,976 个节点，尖峰性能可达到 1 exaFLOPS (1,000 petaFLOPS)。富岳除了在 Linpack 中拿到好成绩也在 HPL-AI 中获得 1.421 exaFLOPS。

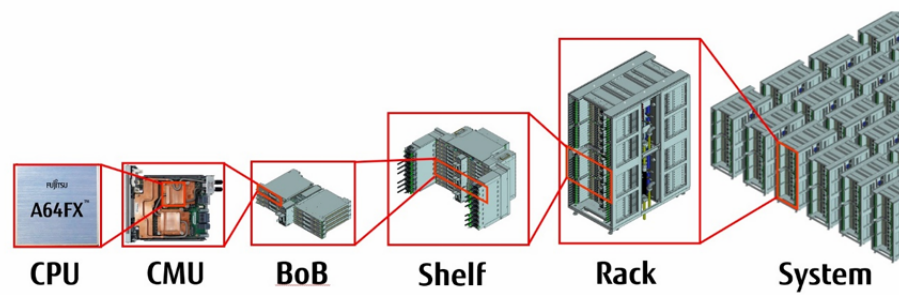


图 4.9: 日本富岳计算机架构 [4]

各指标	富岳	京
运算速度	442 PFLOPS (约 44.2 京次/秒)	10.51 PFLOPS (约 1 京次/秒)
处理器	基于 ARM v8.2-A 指令集的 富士通 A64FX 处理器， 48 核/芯片，台积电 7nm 制程	基于 SPARC64 VIIIfx 架构的 富士通处理器， 8 核/芯片，主频 2.0GHz
节点规模	400 台机柜	800 台机柜
能耗	30MW-40MW	12.7 MW
性能功耗比	16.9 GFLOPS/W	0.83 GFLOPS/W

表 1: 日本富岳与京超算性能对比

如表格1所示，富岳相较于京，性能得到大幅提升。其 ARM 架构也是最为亮点的一个方面。

4.3 Frontier

Frontier 是全球首台实现百亿亿次 (Exascale) 计算的超算，峰值性能达 1.102 exaFLOPS (每秒 1.102E10¹⁸ 次浮点运算)，相当于 10 万台笔记本电脑同时运行的速度。它包含约 5 万个处理器，其中 3.8 万个为 AMD Instinct MI250X GPU，搭配 AMD EPYC CPU，基于惠普企业 (HPE) 的 Cray EX 平台构建。Frontier 采用室温水冷却技术，能耗效率显著提升，冷却能耗仅占总功耗的 3%-4%，相比前代超算 Summit (冷却能耗占比 10%) 效率提高 4 倍。以每瓦 52.23 千兆浮点运算的能效比，曾位列 Green500 榜单首位，成为“最快且最节能”的超算。其优越性通常用于训练大模型。

5 国内外超算技术对比

5.1 重要数据一览

指标	神威·太湖之光	富岳	Frontier	EI Capitan
峰值计算能力	125.4 PFLOPS	1 EFLOPS	1.5 EFLOPS	2+ EFLOPS
能效比	6.05 GFLOPS/W	15.4 GFLOPS/W	52.23 GFLOPS/W	约 55 GFLOPS/W
芯片架构	申威 26010 处理器 MIPS-like 的 64 位申威指令集	富士通 A64FX 处理器 基于 ARM 架构	AMD EPYC 处理器 AMD Radeon GPU	基于 Cray EX Shasta 架构 采用 44,544 个 AMD-MI300A 集成处理器
内核总数	10,649,600 个	7,630,848 个	8,730,112 个	>11,000,000 个
内存容量	1.31PB	4.85PB	4.6PB	NA
存储系统	20PB	150PB	700PB	NA
IO 吞吐量	200GB/s	5TB/s	75TB/s	NA
互联网络	定制的高性能互连网络	富士通自研 Tofu D 互连	HPE Slingshot 高速互连	定制高速互连网络
操作系统	基于 Linux 的 Sunway RaiseOS	基于 Linux 的 富士通定制系统	HPE Cray OS	基于定制操作系统

“数据来源 1 和 2：神威·太湖之光技术白皮书（2016）；富岳 RIKEN 技术报告（2021）”
“Frontier ORNL 官方文档（2022）；EI Capitan LLNL 简报（2024）”

表 2: 中日美三国超算数据对比

表格2向我们展现了中日美三国代表性的超算，中国目前的高端超算没有数据流出。

5.2 数据的分析

如表格2所示，神威·太湖之光在存储容量、能效比等指标上存在显著差距，但其基于自主众核架构的全栈可控性具有战略意义。这种技术路线选择体现了我国对产业链安全的重视——通过牺牲短期能效换取技术主权，规避日本"京"因依赖 SPARC 架构导致的生态封闭风险。

当前国际超算发展呈现两大趋势：一是**存储与内存容量指数级增长**（如 Frontier 存储达 700PB），二是**低延迟互连网络**成为性能瓶颈突破口。我国需针对性提升存储能力并研发自主高速互连技术（如神威互连网络的亚微秒级延迟优化），以支撑气候模拟、基因组学等数据密集型应用。

6 对我国超算技术的展望

未来，我国的超级计算机应该向着轻量化，高速化方向发展。虽然目前我们退出了如 top500 这种超算排行榜，但是我们仍然坚持将量子计算应用到民生领域，让成果落地，让更多人享受红利。同时可以规避“性能竞赛”导致的资源错配。我国应建立自主评价体系，增设“国产化率”“单位 GDP 算力贡献”等指标，引导超算技术向实用化转型 [5]。同时大力培养人才 [6]。

2025 年两会 [7]，全国人大代表郭国平提出构建“四算融合”国家算力体系，整合量子、超算、智算与通算：核心层统筹超算（复杂科学）、智算（AI 训练）、量子集群（算法/加密）、云计算（弹性支持）；边缘层延伸垂直场景；支撑层设专项基金推动跨域研究，促量子企业与国企合作拓展 AI、电网等应用，建混合算力云平台降低中小企业门槛；评估层新增"协同效率"指标（量超交互频次、异构任务率），驱动效能升级。

该体系兼具技术突破与算力主权战略价值。超算技术竞争已从“峰值算力”转向“有效算力”。建议我国建立“场景驱动”研发机制：在航空航天领域侧重低延迟互连技术，在生物医药领域强化分子动力学模拟专用加速器研发，避免陷入“为超算而超算”的误区。这也是国家重点解决的任务。

参考文献

- [1] Wiki. 超级计算机. [EB/OL], [2025-02-20].
- [2] HDS. 共享内存模型 uma numa coma. [EB/OL], [2020-07-07].
- [3] Top 500. November 2024. [EB/OL], [2024-11].
- [4] Jack Dongarra. Report on the fujitsu fugaku system. Technical report, Tech Report No ICL-UT-20-06, University of Tennessee, Innovative Computing ..., 2020.
- [5] 苏诺雅. 中国超算技术赶超发展模式探析. 国防科技大学学报, 43(3):86, 06 2021.
- [6] 司宏伟. 中国超级计算机研制反思从第一台国产超级计算机"银河-"说起. 科学文化评论, 18(1):109–119, 2 2021.
- [7] 中国科学报. 郭国平代表：率先布局“四算融合”的算力体系. [EB/OL], [2025-03-11].