



南開大學
Nankai University

南 開 大 學

計 算 機 學 院
并行政程序设计期末实验报告

并行体系结构调研

周宝亮

年级：2022 级

专业：计算机科学与技术

指导教师：王刚

2025 年 3 月 20 日

摘要

随着计算机技术的迅猛发展，并行计算在提升计算能力方面的重要性日益凸显。并行计算依赖于并行计算机的架构设计以及并行编程的能力。超级计算机通过采用先进的并行架构和庞大的处理器与计算核心数量，实现了高速运算的目标。

本文以并行计算为切入点，系统介绍了并行计算和超级计算机的基本概念。具体分析了当前全球超算排名榜首的“富岳”以及我国自主研发的超级计算机“神威·太湖之光”的体系结构，并进行了对比分析。此外，本文还介绍了这两台超级计算机的并行编程技术和相关软件，并对超级计算机未来可能的发展方向进行了探讨。

关键字：并行计算、超级计算机、体系结构、富岳、神威·太湖之光 s

目录

一、 并行计算简介 1

 (一) 并行计算 1

 (二) 并行计算机 1

 (三) 并行编程 1

二、 超级计算机简介 1

 (一) 超级计算机的定义与特点 1

 (二) 发展历程与性能指标 2

 (三) 典型应用场景 2

 (四) 中国超级计算机发展概况 2

三、 具体超算架构分析 2

 (一) 富岳 2

 1. 节点组成 2

 2. 处理器架构 3

 3. 性能 3

 4. 网络 4

 5. 数据 4

 (二) 神威·太湖之光 5

 (三) 对比分析 6

四、 软件生态分析 7

 (一) 富岳 7

 (二) 神威·太湖之光 7

五、 总结 7

一、 并行计算简介

(一) 并行计算

并行计算（英语：parallel computing）通常是指多个指令能够同时执行的一种计算模式。在实现同时执行的前提下，可以将计算过程分解为较小的部分，然后通过并发的方式来解决问題。

并行计算可以分为时间并行和空间并行。时间并行例如指令的流水线运行，而空间并行则是指多个处理器协同执行并发计算。此外，根据任务的组织方式，还可以将并行计算分为数据并行和任务并行。在数据并行中，大的任务会被拆解为若干个子任务；在任务并行中，每一个线程负责执行一个分配到的任务，这些线程会被分配到该并行计算体系中的各个计算节点上进行处理。

(二) 并行计算机

根据 Flynn 分类法，计算机系统可以分为四种结构：单指令流单数据流（SISD）、单指令流多数据流（SIMD）、多指令流单数据流（MISD）、多指令流多数据流（MIMD）。其中，SIMD 和 MIMD 是并行计算机的两类基本结构。MIMD 类型的机器又可以进一步细分为五种常见类型：并行向量处理机（PVP）、对称多处理机（SMP）、大规模并行处理机（MPP）、工作站机群（COW）、分布式共享存储机（DSM）。

并行计算机至少包含两台或两台以上的处理器，这些处理器通过互连网络相互连接，实现通信。一般来说，互连网络的连接方式可以分为静态连接，例如一维线性连接、超立方体连接、树连接等，以及动态连接，例如总线连接（Bus）、交叉开关（CS）、多级互连网络（MIN）等。

并行计算机有五种访存模型：均匀访存模型（UMA）、非均匀访存模型（NUMA）、全高速缓存访存模型（COMA）、一致性高速缓存非均匀存储访问模型（CC-NUMA）、非远程存储访问模型（NORMA）。

(三) 并行编程

在现代高性能计算体系中，并行计算的实现绝非仅靠并行计算机硬件就能完成，它需要并行编程技术与并行软件体系的深度协同支撑。从技术本质来看，并行编程的核心价值在于打破单处理器的性能局限，通过任务拆解与分配，让多个计算单元同步作业，从而大幅提升复杂运算效率，满足科学计算、大数据处理等场景的高算力需求。程序员需要依托并行机提供的编程环境——无论是分布式内存架构下的消息传递模型，还是共享内存架构下的多线程编程模型——将并行算法转化为可执行的程序代码。这一过程涉及对计算任务的精细划分，确保多核处理器或多计算节点能够高效协作，实现数据处理与运算的并行化。

当前并行编程领域存在显著的技术特性：由于不同并行机的体系结构差异巨大，如集群计算系统、异构计算平台（CPU+GPU）、超级计算机等，其硬件组织形式、数据交互机制各有不同。这种多样性导致统一的并行编程软件与范式难以形成。开发者需根据具体硬件架构，选择适配的编程模型，例如用于分布式内存的 MPI（消息传递接口）、支持共享内存多线程的 OpenMP，或是针对 GPU 计算的 CUDA 等。

二、 超级计算机简介

(一) 超级计算机的定义与特点

超级计算机（Supercomputer）是指具备超高计算能力、超大存储容量和高速数据传输能力的计算机系统，其核心特征在于通过大规模并行处理技术实现每秒万亿次（TeraFLOPS）乃至亿

亿次 (PetaFLOPS) 的浮点运算能力。与普通计算机相比, 超级计算机在处理器数量 (可达数万至数百万核心)、内存带宽、网络延迟及存储系统吞吐量等方面具有数量级优势, 专为解决复杂科学计算、大规模数据模拟等传统计算机无法胜任的任务而设计。

(二) 发展历程与性能指标

自 1964 年 CDC 6600 问世以来, 超级计算机经历了从向量机 (如 Cray-1) 到大规模并行处理 (MPP) 系统的演变。现代超级计算机的性能通常以每秒浮点运算次数 (FLOPS) 衡量, 其排名依据 LINPACK 基准测试结果定期更新于 TOP500 榜单。当前顶尖超算的峰值性能已突破 ExaFLOPS (每秒百亿亿次运算) 量级, 标志着超级计算机正式迈入 “E 级计算” 时代。

(三) 典型应用场景

作为国家重要的科技支柱, 超级计算机在许多关键领域都发挥着独一无二的作用。在气候研究方面, 它能通过处理海量数据建立地球气候模型, 帮助科学家预测未来几十年的气候变化趋势。在能源开发领域, 超级计算机可以模拟极端条件下的物质状态, 为核聚变发电装置的设计提供关键支持。在生物医学研究中, 它能快速模拟病毒结构和药物分子的作用过程, 大幅缩短新药研发时间。在国家安全方面, 超级计算机不仅能进行武器系统的虚拟测试, 还能模拟复杂环境下的物理现象, 为重大工程的设计提供可靠依据。

(四) 中国超级计算机发展概况

我国通过 “天河” “神威” 系列实现技术突破, 其中神威·太湖之光 (峰值性能 125.4 PFLOPS) 曾四次蝉联 TOP500 榜首。其采用全国产申威处理器与神威网络架构, 软件生态覆盖气候、海洋、材料等 20 余个领域, 标志着我国在超算硬件设计与应用开发方面达到国际领先水平。

三、具体超算架构分析

(一) 富岳

根据 2024 年 11 月的 TOP500 排名统计, 由日本神户理研计算科学中心 (R-CCS) 与富士通共同开发的超级计算机 “富岳” (FUGAKU) 排名第六。这台基于富士通定制 ARM A64FX 处理器的超级计算机, 拥有高达 763,0848 个内核, 其性能表现尤为突出, 实现了 442 Pflop/s 的 HPL 基准测试得分,

1. 节点组成

富岳 (Fugaku) 超级计算机的节点总数为 158,976。其架构设计如下:

- 单个 CPU 组成一个节点, 两个 CPU (即两个节点) 被安装在一个称为 **CPU 内存单元 (CMU)** 的板上。
- 八个 CMU 组成一个 “刀片束 (BoB)”, 这意味着每个 BoB 包含 16 个节点。
- 三个 BoB 组成一个 Shelf, 因此每个 Shelf 包含 48 个节点。
- 8 个 Shelf (共 384 个节点) 安装在一个计算机机架中, 但需要注意的是, 有些机架可能只容纳 192 个节点。

整个系统由 **432** 个机架组成，其中：

- **396** 个机架每台包含 384 个节点，
- 剩余的 **36** 个机架每台包含 192 个节点。

通过这种设计，富岳最终实现了总计 **158,976** 个节点的规模。

2. 处理器架构

富岳超级计算机采用了 Armv8.2-A SVE（可伸缩向量扩展）指令集架构，支持 512 位 SIMD（单指令多数据）。它是全球首台采用 ARM 架构并荣登 TOP500 榜首的超算。这一成就得益于富士通与 ARM 的合作，共同推出了 SVE 指令集。通过为每个计算核心配备双流水线设计，SVE 能够高效支持 512 位 SIMD，从而显著提升浮点运算性能。这种架构的创新极大地增强了富岳在高性能计算领域的竞争力。

富岳的每个计算节点配备了一颗 A64FX CPU。这颗 CPU 由 48 个计算核心和 4 个辅助核心组成，被划分为四个 CMG（Core Memory Group）。每个 CMG 包含 13 个核心、一个 L2 缓存以及一个内存控制器。四个 CMG 通过 ccNUMA（Cache-Coherent Non-Uniform Memory Architecture）架构和片上目录机制保持缓存一致性。此外，CMG 内部通过 X-bar 结构连接，显著提升了 L2 缓存的吞吐量。同时，CMG 中的进程绑定支持高达 48 个核心的线性扩展能力，从而实现高效的并行处理。具体结构如图 1 所示。

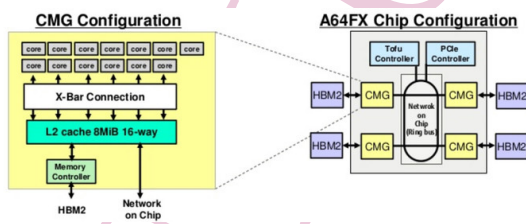


图 1: 结构示意图

A64FX 采用了高性能的 HBM2 内存，每个 CMG（Core Memory Group）单元配备 8GB HBM2，单个 CMG 的带宽达到 256GB/s。整个系统共配置了 32GB HBM2，总带宽高达 1024GB/s。这种超高带宽的设计是富岳性能的一大亮点。此外，A64FX 在核心（cores）、缓存（caches）和内存控制器（memory controllers）中引入了无序机制，这一机制能够最大化每一层的带宽能力，从而显著提升整体性能。

3. 性能

在峰值理论性能方面，富岳支持两种运行模式：normal 模式和 boost 模式。在 normal 模式下，CPU 的时钟频率为 2 GHz，此时的理论峰值性能如下：

- 双精度浮点性能（64 位）：488 PFlops；
- 单精度浮点性能（32 位）：977 PFlops；
- 半精度浮点性能（AI 学习，16 位）：1.95 EFlops；
- 整数性能（AI 推理，8 位）：3.90 EOps。

而在 boost 模式下，CPU 的时钟频率提升至 2.2 GHz，理论峰值性能进一步提高：

- 双精度浮点性能 (64 位): 537 PFlops;
- 单精度浮点性能 (32 位): 1.07 EFlops;
- 半精度浮点性能 (AI 学习, 16 位): 2.15 EFlops;
- 整数性能 (AI 推理, 8 位): 4.30 EOps。

在浮点运算性能方面, A64FX 的每个核心配备了双流水线 SVE 512 位 SIMD。每个 SIMD 单元能够同时执行两条 FMA (fused multiply-add) 指令, 因此单核每周期可提供

$$\frac{2 \times 512 \text{ 位} \times 2 \text{ 条 FMA 指令}}{64 \text{ 位}} = 32 \text{ FLOPS}$$

的双精度浮点性能。如果 A64FX 以 2.2 GHz 的频率运行, 则理论上每个 CPU 最高可提供

$$32 \text{ FLOPS} \times 2.2 \text{ GHz} \times 48 \text{ cores} = 3379.2 \text{ GFlop/s}$$

的双精度浮点性能。

4. 网络

富岳的一大亮点在于其创新的互联技术, 它采用了富士通自主研发的 Tofu Interconnect 系列中的 Tofu Interconnect D (简称 TofuD)。其中, “Tofu” 源自 “Torus Fusion”, 象征着网络结构的融合; 而 “D” 则代表 “Density” 和 “Dynamic”, 体现了高节点密度、动态分组切片以及网络故障恢复能力的特点。

TofuD 在设计上具有更先进的评估机制, 能够实现更低的延迟, 并增加了 Tofu Barrier 资源。这种设计在追求高吞吐量 (Put Throughputs) 和注入率 (Injection Rates) 的同时, 依然保持了高效的性能表现。

在互联结构方面, 每个计算节点配备了六个 Tofu 网络接口 (TNI), 每个接口的传输速度为 6.8 GB/s, 总带宽达到 40.8 GB/s。这些接口连接到一个具有 10 个端口的 Tofu 网络路由器, 每个端口包含 2 个通道, 总共提供 20 个通道。如图 5 所示。

此外, 大约有 200,000 根 Tofu 电缆用于连接 Fugaku 的 CPU。其中, 约一半的光缆在机架之间运行, 总长度约为 900 公里。

5. 数据

	Cores	Memory	Processor	Interconnect
富岳	7630848	5087232GB	A64FX 48C 2.2GHz	Tofu interconnect D

表 1: 硬件

	Linpack Performance (Rmax)	Theoretical Peak (Rpeak)	Nmax	HPCG [TFlop/s]
富岳	442010 TFlop/s	537212 TFlop/s	21288960	16004.5

表 2: 性能

超级计算机	Power	Power Measurement Level
富岳	29,899.23 kW (Optimized: 26248.36 kW)	2

表 3: 功耗

(二) 神威·太湖之光

神威·太湖之光 (Sunway TaihuLight) 是由中国国家并行计算机工程与技术研究中心 (NR-CPC) 开发, 并部署在中国江苏省无锡市国家超级计算中心的高性能计算系统。该系统在 2021 年 11 月的 TOP500 排名中位列第四, 峰值性能达到 93 PFlop/s。(PS: 查看最新数据非前十需要登录下载, 所以这里使用了旧数据)

神威·太湖之光采用众核 + CPU 的混合架构, 核心处理器为我国自主研发的 SW26010 处理器。每颗 SW26010 处理器集成了多达 260 个核心 (每四个核心组成一个处理单元), 主频为 1.45 GHz, 配备 32 GB 内存。整套系统共使用了 40960 颗处理器, 总计拥有 10649600 个核心和 1310720 GB 的内存容量。其理论峰值性能高达 125435.9 TFlops, 实际最大性能为 93014.6 TFlops, 计算效率达到 74%。



图 2: 总体结构

处理器芯片的架构设计如下: 芯片由 4 个核心组 (Core Groups, CGs) 构成, 各核心组通过片上网络 (NoC) 实现互联。每个核心组包含一个管理处理单元 (Management Processing Element, MPE) 和 64 个计算处理单元 (Computing Processing Elements, CPEs), 这些单元以 8×8 的网格形式排列。此外, 每个核心组拥有独立的内存空间, 并通过内存控制器 (MC) 与 MPE 和 CPE 集群相连。整个处理器通过系统接口 (System Interface, SI) 与其他外部设备进行通信, 其整体结构如图 3 所示。

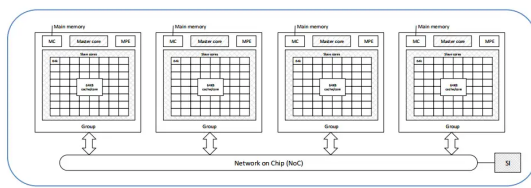


图 3: 节点架构

(三) 对比分析

神威·太湖之光与富岳的对比数据如表 4、表 5、表 6 所示。

超级计算机	Cores	Memory	Processor	Interconnect
富岳	7630848	5087232 GB	A64FX 48C 2.2GHz	Tofu interconnect D
神威·太湖之光	10649600	1310720 GB	Sunway SW26010 260C 1.45GHz	Sunway

表 4: 硬件数据

超级计算机	Linpack Performance (Rmax)	Theoretical Peak (Rpeak)	Nmax	HPCG [TFlop/s]
富岳	442010 TFlop/s	537212 TFlop/s	21288960	16004.5
神威·太湖之光	93014.6 TFlop/s	125436 TFlop/s	12288000	480.848

表 5: 性能

超级计算机	Power	Power Measurement Level
富岳	29,899.23 kW (Optimized: 26248.36 kW)	2
神威·太湖之光	15371.00 kW (Submitted)	2

表 6: 功耗

从数据层面来看，相较于富岳，神威·太湖之光在内存容量、CPU 主频、运算速度以及峰值理论性能等方面略显逊色。

从架构角度来看，神威·太湖之光的核心数量多于富岳。然而，单核性能方面，神威的表现相对较弱，其 28nm 工艺的 2GHz 核心性能大致相当于 ARM A75-A76 的水平，与 ARM 阵营中苹果 A12、A13 所采用的 CPU 核心结构相比存在较大差距。不过，神威的强大之处在于其先进的架构设计，核心之间的通信开销较小，这表明中国在工艺和单核性能提升方面仍有进步空间。同时，SW26010 处理器采用了改进型 Alpha 指令集，基于 64 位 RISC 架构，专为高性能计算而设计。神威·太湖之光的简单架构正是其运算速度超越其他高能耗 HPC 系统的关键所在。然而，神威芯片使用定制的 64 位指令集，频率仅为中等水平（1.45GHz），且每个核心仅能执行一个线程（不支持超线程），软件生态也远不如 Intel 丰富。

四、 软件生态分析

并行计算作为现代高性能计算的核心技术，其发展离不开硬件与软件的协同作用。硬件是并行计算的基础，为大规模数据处理和复杂任务提供了强大的算力支持；然而，仅仅依靠硬件的强大性能远远不够，还需要配套的软件系统来充分发挥硬件的优势。只有通过精心设计的并行算法、高效的调度机制以及优化的编程模型，才能将硬件的潜力最大化，从而实现高效的计算。

（一） 富岳

富岳使用了 FUJITSU Software Technical Computing Suite V4.0 编译器，该编译器支持多种编程语言和扩展，包括 Fortran 2008 和 Fortran 2018、带有 GNU 和 Clang 扩展的 C11、带有 GNU 和 Clang 扩展的 C++14 和 C++17、OpenMP 4.5 和 OpenMP 5.0，以及 Java。

在并行编程方面，富岳采用了 XcalableMP 和 FDPS。其中，XcalableMP (XMP) 是由 XMP SpecWG 提出的一种面向分布式内存的 PGAS (Partitioned Global Address Space) 编程模型和语言，旨在提升并行编程的生产力和性能。XMP 规范的 1.4 版本引入了多项新特性，例如融合了 OpenMP 和 OpenACC 的功能，并提供了用于集体通信的库。这种语言结合了全局视图编程与全局视图分布式数据结构的数据并行性，同时具备基于指令的 Fortran 和 C 的 PGAS 模型的语言扩展能力。

FDPS (Framework for Developing Particle Simulator) 是一个专门用于开发粒子模拟器的框架。FDPS 的 API 提供了动态负载平衡、节点间通信以及力的计算等功能，为大规模粒子模拟任务提供了高效的支持。

此外，富岳运行在 Red Hat Enterprise Linux 操作系统之上。

（二） 神威·太湖之光

神威·太湖之光采用 Raise Linux 操作系统，支持多种编程语言，包括 C、C++ 和 Fortran，并且兼容 MPI (Message Passing Interface)、OpenMP (Open Multi-Processing) 和 OpenACC (Open Accelerators for Computing) 等主流并行编程模型与环境。

五、 总结

本文以并行计算的概念为切入点，引出对超级计算机的介绍。文章详细分析了富岳、神威·太湖之光等超级计算机的体系结构，并进行了对比分析。此外，还简要探讨了超算的并行编程与软件技术。最后，文章展望了超算的发展趋势，分析了未来可能的发展方向。

参考文献

NIKU