



南开大学  
Nankai University

南 开 大 学

计 算 机 学 院

并行程序设计实验报告

---

## 并行计算机体系结构调研

---

姓名：韩佳迅

学号：2012682

年级：2020 级

专业：计算机科学与技术

指导教师：王刚

2022 年 3 月 4 日

## 摘要

随着计算机的飞速发展，目前**并行计算**在提升计算能力上越来越重要，并行计算离不开并行计算机的架构和并行编程的能力。超级计算机采用优秀的并行架构和巨大的处理器和计算核心数目，达到高速运算的结果。

本文通过并行计算引入，分析介绍了并行计算、**超级计算机**的概念。具体分析了当前 top500 榜首的**富岳**和全国产的超算 **神威·太湖之光**的**体系结构**，并进行对比分析。除此之外，介绍两台超算的并行编程与软件，并就未来超级计算机的可能发展方向进行分析。

**关键字：**并行计算、超级计算机、富岳、神威·太湖之光、体系结构

## 目录

<b>一、 并行计算简介</b>	<b>1</b>
(一) 并行计算 . . . . .	1
(二) 并行计算机 . . . . .	1
(三) 并行编程 . . . . .	1
<b>二、 超级计算机简介</b>	<b>1</b>
(一) 超级计算机 . . . . .	1
(二) 超级计算机的发展 . . . . .	2
<b>三、 具体超级计算机的架构分析</b>	<b>2</b>
(一) 富岳 . . . . .	2
1. 节点组成 . . . . .	2
2. 处理器及架构 . . . . .	2
3. 性能 . . . . .	3
4. 网络 . . . . .	3
(二) 神威·太湖之光 . . . . .	4
(三) 对比分析 . . . . .	5
<b>四、 编程及软件分析</b>	<b>6</b>
(一) 富岳 . . . . .	6
(二) 神威·太湖之光 . . . . .	6
<b>五、 未来发展趋势</b>	<b>7</b>
<b>六、 总结</b>	<b>7</b>

## 一、 并行计算简介

### (一) 并行计算

并行计算（英语：parallel computing）一般是指许多指令得以同时进行的计算模式。在同时进行的前提下，可以将计算的过程分解成小部分，之后以并发方式来加以解决。[3]

并行计算可以分为时间并行和空间并行。时间并行如指令的流水线运行，空间并行则是指多个处理器执行并发计算。同时，并行计算还可以分为数据并行和任务并行，在数据并行中，大的任务化解成若干个子任务；在任务并行中，每一个线程执行一个分配到的任务，而这些线程则被分配到该并行计算体系的各个计算节点中去。

### (二) 并行计算机

根据 Flynn 分类法，计算机系统可分为：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）四种结构。其中，SIMD 和 MIMD 即为并行计算机的两类基本结构。MIMD 类的机器又可分为常见的五类：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）、分布式共享存储处理机（DSM）。

并行计算机至少包含两台或两台以上处理器，这些处理器通过互连网络相互连接，相互通信。一般来说，分为静态连接，如一维线性连接，超立方体连接，树连接等，和动态连接，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

要想实现并行计算，只有并行计算机是无法完成的，还需并行编程、并行软件的支持。程序员需要在并行机提供的并行编程环境上，具体实现并行算法，编制并行程序。由于不同并行机的体系结构不同，目前还没有统一的并行编程软件及编程方式。

## 二、 超级计算机简介

### (一) 超级计算机

超级计算机（英语：Supercomputer），指能够执行一般个人电脑无法处理的高速运算的计算机，规格与性能比个人计算机强大许多。超级计算机也是并行计算机的范畴。

超级计算机，由超多个计算节点组成，每个计算节点配有 CPU、GPU 以及专用处理器，节点之间由高速网络互联。如图1所示。

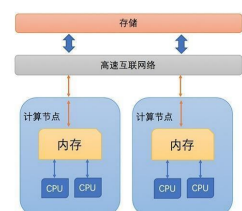


图 1: 超级计算机硬件示意图

超级计算机使用调度器分配计算资源，服务方先提供好一批计算资源放置在那里，使用方如有需求，向调度器申请，如有闲置的资源则分配给需求方。

从技术角度说，超级计算机可以分为三层：基础层、中间层与应用层。基础层是超级计算机的核心，它包括了超级计算机的所有硬件设备、架构、电源供应系统、冷却系统，以及大量辅助设备组成的大型并行运算系统。中间层的主要工作是为应用层提供技术、数据与算法支持，它包括六大设备与三大网络，负责用户登录、任务运算、数据存储等。应用层一般指人们利用超级计算机搭建的各种任务系统，比如地球模拟、天气模拟等。[5]

## （二） 超级计算机的发展

超级计算机的性能用“每秒浮点运算次数”（FLOPS）来表示，超级计算机的性能发展近似符合摩尔定律，算力以倍数增长，如图2所示。

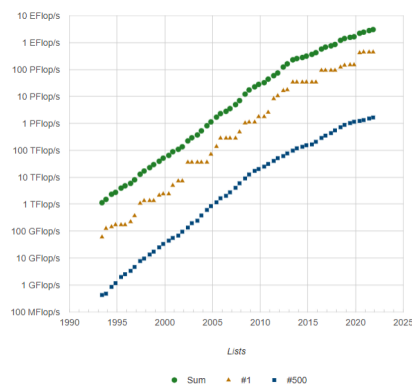


图 2

## 三、 具体超级计算机的架构分析

### （一） 富岳

根据 2021 年 11 月 top500 的结果统计，安装在日本神户的理研计算科学中心 (R-CCS)，由理研和富士通共同开发，基于富士通的定制 ARM A64FX 处理器的超级计算机富岳 (FUGAKU) 位于榜首，它有 7630848 个内核，使其能够实现 442 Pflop/s 的 HPL 基准测试得分，领先位于第二的 Summit 3 倍。

#### 1. 节点组成

富岳中的节点总数为 158976。单个 CPU 组成一个节点，两个 CPU（两个节点）安装在称为 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 有 16 个节点。三个 BoB 组成一个 Shelf，因此每个 Shelf 有 48 个节点。8 个 Shelf (384 个节点) 安装在一个计算机机架中（有些机架有 192 个节点）。Fugaku 由 432 个机架组成，其中 396 个机架有 384 个节点，其余 36 个机架有 192 个节点。这使得总共有 158,976 个节点。

#### 2. 处理器及架构

富岳采用 Armv8.2-A SVE 512 位的指令集架构，是首台获得 TOP500 榜首的 ARM 架构超算。富士通并联合 ARM 推出了 SVE（可伸缩矢量扩展）指令集，每个计算核心都拥有双流水线支持 SVE 的 512 位 SIMD，大幅强化浮点性能。

富岳的每个计算节点包含一个 A64FX CPU。A64FX 有 48 个计算核心及 4 个辅助核心组成，分为四个 CMG (Core Memory Group)，每个 CMG 由 13 个核、一个 L2cache 和一个 memory controllers 组成。四个 CMG 通过 ccNUMA 与 on-chip directory 保持 cache 一致性，CMG 中的 X-bar 连接提高 L2cache 的吞吐量，CMG 中的进程绑定允许多达 48 个核的线性可伸缩性。[1] 如图3所示。

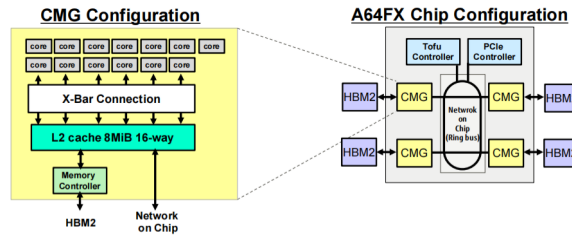


图 3: 多核架构

A64FX 使用高性能的 HBM 2 内存，每个 CMG 单元配备 8GB HBM2，带宽 256GB/s，总计 32GB HBM2，1024GB/s 带宽。caches 和 memory 中的超高带宽是富岳的一大特点。A64FX 在 cores、caches 和 memory controllers 中存在无序机制，它使每一层的带宽能力最大化。[1] 如图4所示。

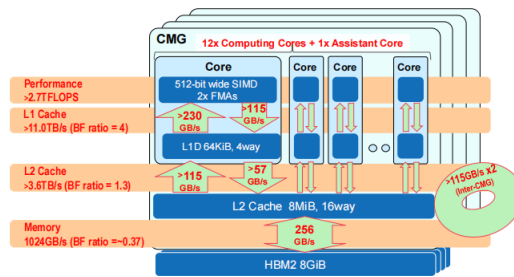


图 4

### 3. 性能

在峰值理论性能方面，富岳可以在 normal 模式（CPU 时钟速度为 2 GHz）和 boost 模式（CPU 时钟速度为 2.2 GHz）下运行。在 normal 模式下，CPU 时钟速度达到 2GHz，理论峰值双精度性能（64 位）达到 488 PFlops，理论峰值单精度性能（32 位）977 PFlops，理论峰值半精度（AI 学习）性能（16 位）1.95 EFlops，理论峰值整数（AI 推理）性能（8 位）3.90 EOps。在 Boost 模式下，CPU 时钟速度可达 2.2 GHz，理论峰值双精度性能（64 位）537 PFlops，理论峰值单精度性能（32 位）1.07 EFlops，理论峰值半精度（AI 学习）性能（16 位）2.15 EFlops，理论峰值整数（AI 推理）性能（8bit）4.30 EOps。

在浮点运算性能方面，A64FX 每核拥有双流水线 SVE 512 位 SIMD，而每个 SIMD 可以同时执行两条 FMA 指令，因此单核每周期可提供  $2 \text{ pipelines} * 512 \text{ bit} * 2 \text{ FMAs} / 64 \text{ bit} = 32 \text{ FLOPS}$  的双精度浮点性能。若 A64FX 以 2.2 GHz 的频率运行，则理论上每个 CPU 最高可提供  $32 \text{ FLOPS} * 2.2 \text{ GHz} * 48 \text{ cores} = 3379.2 \text{ GFlop/s}$  双精度浮点性能。

### 4. 网络

富岳的一大亮点在于互联方式的创新，它采用了富士通自研 Tofu Interconnect 系列中的 Tofu Interconnect D(TofuD)，其中 Tofu 代表“Torus Fusion”，D 代表“Density”和“Dynamic”，

意为高节点密度、动态分组切片及其带来的网络故障恢复能力。TofuD 具有更先进的评估环境、更低的延迟,并增加了 Tofu Barrier 资源;在注重吞吐量 (Put Throughputs) 和注入率 (Injection Rates) 的同时也保持高效率。

互联结构有六个 Tofu 网络接口 (TNI, 每个执行速度为 6.8 GB/s), 总共 40.8 GB/s, 连接到具有 10 个端口的 Tofu 网络路由器, 每个端口有 2 个通道 (总共 20 个通道)。如图5所示。并且, 大约 200000 根 Tofu 电缆连接 Fugaku 的 CPU。其中大约一半是在机架之间运行的光缆, 总长度约为 900 公里。

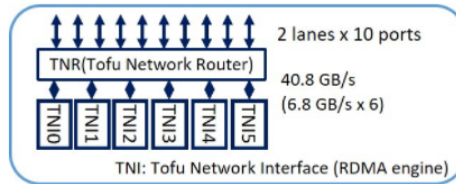


图 5: TofuD 互联结构

富岳的相关数据 [2] 如表1、表2、表3所示。

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

表 1: 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

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) 开发并安装在中国江苏省无锡市国家超级计算中心的系统, 以 93 Pflop/s 在 top500 2021 年 11 月统计列表中排名第四。

整机架构: 神威·太湖之光采用众核 +CPU 架构, 使用我国自主研发的 SW26010 处理器, 每一颗 SW26010 处理器集成多达 260 个核心 (每四个核心一组), 主频 1.45GHz, 搭配 32GB 内存。整套系统使用了 40960 颗处理器, 总计拥有 10649600 颗核心、1310720G 内存, 理论峰值性能 125435.9TFlops, 实际最大性能 93014.6TFlops, 效率达 74%, 如图6。

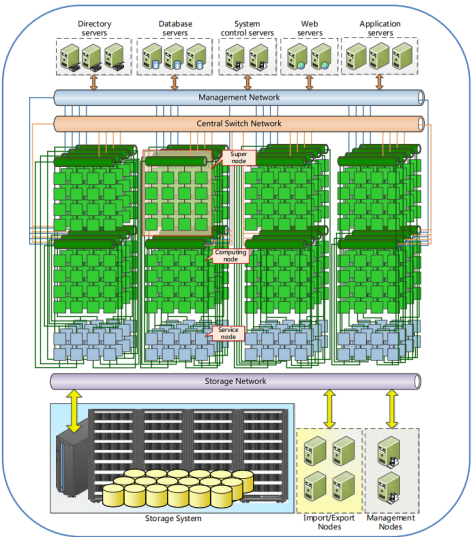


图 6: 总体架构

处理器架构: 处理器芯片由 4 个 core groups(CGs) 组成, 通过 NoC 连接。每个 core groups 包括一个 Management Processing Element(MPE) 和 64 个 Computing Processing Elements (CPEs) 组合在一个 8\*8 的网格中。每个 CG 都有自己的内存空间, 它通过 MC 连接到 MPE 和 CPE 集群。该处理器通过一个系统接口 (SI) 连接到其他外部设备, 结构如图7。[4]

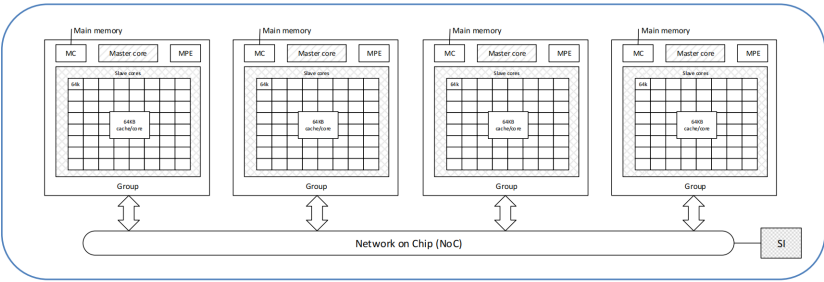


图 7: 节点架构

(三) 对比分析

神威·太湖之光与富岳的对比数据 [2] 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

从数据上来讲, 相较于富岳, 神威太湖之光在内存、CPU 主频、运算速度、峰值理论性能等方面均略逊一筹。

从架构上来讲, 神威·太湖之光的核心数要大于富岳, 看单核心性能, 神威比较来讲很羸弱, 28nm 2Ghz 大约是 2Ghz ARM A75-A76 的水平, 与 ARM 阵营像苹果 A12、A13 使用的 CPU



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

核心结构落后很多。但神威强大在于架构先进，核心之间互联通讯的开销小，这也说明了中国仍需提升工艺与单核心性能。同时，SW26010 采用改进型 alpha 指令集，64 位 RISC（精简指令集）架构，专为高性能计算研发。神威太湖之光专为提升运算速度设计，简单的架构正是它运算速度胜过其他高能耗 HPC 系统的原因。不过，申威芯片采用定制 64 位指令集，频率处于中等水平（1.45GHz），而且每个核心只能执行一个线程（不支持超线程），软件支持也没有 Intel 那么丰富。

从发展趋势上讲，全球仍在等待 Exascale，而从 2017 年神威·太湖之光登顶到 2021 年富岳登顶，超级计算机呈现出高速发展趋势，预计在不久之后，运算性能就可以达到 Exaflop 级别的提升。

## 四、编程及软件分析

### （一）富岳

富岳采用 FUJITSU Software Technical Computing Suite V4.0 编译器，支持 Fortran 2008 and 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 and OpenMP 5.0 以及 java。

富岳的并行编程采用 XcalableMP 和 FDPS。

XcalableMP (XMP) 是由 XMP SpecWG 提出的一种针对分布式内存的 PGAS 编程模型和语言，用于提高并行编程的生产力和性能。XMP 规范 1.4 版本的新特性：混合了 OpenMP 和 OpenACC，具有用于集体通信的库。该语言具有全局视图编程与全局视图分布式数据结构的数据并行性、基于指令的 Fortran 和 C 的 PGAS 模型的语言扩展的特点。FDPS，全称 Framework for Developing Particle Simulator，是框架开发粒子模拟器，FDPS 的 api 提供动态负载平衡、节点间的通信和力的计算。

富岳采用 Red Hat Enterprise Linux 操作系统。

### （二）神威·太湖之光

神威·太湖之光采用 Raise Linux 操作系统，支持 C、C++、Fortran 编程语言，适用于 MPI、OpenMP、OpenACC 等并行语言及环境。

## 五、 未来发展趋势

目前最强的超级计算机“富岳”的运算速度是 442PFLOPS，在机器学习和人工智能应用中经常使用的单精度或进一步降低的精度中，Fugaku 的峰值性能在 1,000 PFlop/s (1 Exaflop/s) 以上。随着超算的发展，不久后的超算将突破 Exaflop 级别。

目前超算的架构虽然发展，但整体变化并不大。相比之下，未来更大规模的改变和创新不是没有可能。区块链网络与分布式计算网络在某种程度上为我们设计超级计算机内部的高速网络与管理系统带来了新的思路和视角，而量子计算机、光子计算机与自旋计算机等却可能更加直接地为超级计算机带来速度上的提升。正在理论发展阶段的光子计算机和自旋计算机分别使用光子代替电子、用电子自旋的磁效应代替电子本身的运动来传递数据，因此将拥有更快的信息传输与处理速度。未来量子计算机、光子计算机和自旋计算机一旦发展成熟，将可能构建出具有全新架构与超凡性能的超级计算机。

## 六、 总结

本文通过并行计算的概念引出对超级计算机的介绍，具体分析了富岳、神威·太湖之光的体系结构并进行了对比分析，还简单分析了超算的并行编程与软件。最后，就超算的发展趋势，分析了未来超算可能的发展方向。

## 参考文献

- [1] Riken 计算科学中心. <https://www.r-ccs.riken.jp/en/>.
- [2] top500. <https://www.top500.org/>.
- [3] **平行計算基礎理論, 系統及應用研究**. 國立中正大學資訊工程研究所, 1992.
- [4] Jack Dongarra. Report on the sunway taihulight system. *PDF*). *www. netlib. org. Retrieved June, 20, 2016.*
- [5] 张磊. 超级计算机到底“超”在哪. **中学科技**, pages 2–7, 2021.