



南開大學
Nankai University

计算机学院

超级计算机并行体系架构调研报告

姓名：王浩

学号：2210207

专业：计算机科学与技术

2024 年 3 月 12 日

目录

1	国际超算并行架构发展历史	2
1.1	20 世纪 60 年代	2
1.2	20 世纪 70 年代	2
1.3	20 世纪 80 年代	2
1.4	20 世纪 90 年代	3
1.5	21 世纪至今	4
2	中国超级计算机发展历史	5
2.1	银河一号	5
2.2	曙光系列	5
2.3	天河系列	5
2.4	神威·太湖之光	5
2.5	现状	6
3	发展超级计算机的意义	6

1 国际超算并行架构发展历史

1.1 20 世纪 60 年代

1961 年, IBM 研制出 IBM 7030 Stretch 计算机, 是世界上第一台通用的指令流水线化 (Pipeline) 电脑, 利用四级流水线, 互相重叠指令提取、解码和执行得到 1.6 倍效能, 能够在 24h 内执行超过 300 亿次乘法运算, 最多可同时运行 9 个程序。1964 年, 西蒙克雷成功研制出巨型计算机 CDC6600, 运算速度达 1MFLOPs, 是 IBM 7030 的三倍。其首次运用了固定交错法 (Fixed Interleave) 来实现多线程。如图所示, 假设有 N 个线程, 每个线程都能在 N 个周期内执行一条指令, 如果一个线程的指令没有执行完成, 则在后续线程中插入空闲周期 (Bubble), 以便等待其执行完成。但是这种方法带来的性能提升不够显著, 因为线程之间的工作量分配不均衡, 经常需要插入空闲周期。



图 1.1: Fixed Interleave

1.2 20 世纪 70 年代

1971 年, TI ASC PPUs 首次使用了基于软件控制的交错法 (Software-Controlled Interleave), 如图所示, 由操作系统显示控制线程交错运行, 图中蓝色线程的调度频率大约是橙色和紫色的两倍。这种方式比固定交错法能提升更多的性能, 但是会给操作系统带来更多的开销。1972 年, 世界第一台并行计算机 ILLIAC IV 诞生, 可归类为 SIMD (单指令流多数据流) 型计算机。ILLIAC 的中央处理装置分成四个可以执行单独指令组的控制器, 每个控制器管理 64 个处理单元。每个处理单元拥有 2048 字存片器, 可以单独作为一个运算装置, 并能和所有其它的处理单元发生联系。同年代诞生的 SIMD 并行计算机还有阵列机 ICLDAP、Good-year MPP, 向量机 CRAY-1、STAR-100 等。



图 1.2: Software-Controlled Interleave

1.3 20 世纪 80 年代

1982 年诞生的 Denelcor HEP 携带 8 个时钟频率为 10MHz 的处理器, 每个处理器支持 120 个线程, 可归类为 MIMD (多指令流多数据流)。它是第一台商业化并且使用基于硬件控制的调度法 (Hardware-Controlled Thread Scheduling) 的机器, 如图所示, 其线程调度是通过处理器的硬件部分负责确定何时执行哪个线程, 并且在执行线程时可在不同的线程之间进行切换。与操作系统调度相比, 该方法具有更低的延迟和更高的吞吐量。在 80 年代中期, 共享存储多处理机系统得到了稳定发展, 两个成功的机器为 Sequential (20 个结点)、Encore (16-32 个结点)。与此同时, 人们对这种系统的内存访问瓶颈有了清楚认识, 开始寻求解决办法。在此期间还诞生了 MIMD 架构的 MPPu CUBE 超级计算机, 含有 1024 个结点, 每个结点内含有 CPU 和存储单元, 所有结点通过超立方体网络相互连接, 以便进行数据交换和协作计算。80 年代后期, 真正具有强大并行计算能力的计算机开始出现: SIMD 架构的 CM-2 有 32000 个处理单元, 对 Linpack 测试获得了 5.2GFLOPs 的性能; 超立方体连接的分布存储 MIMD 架构并行机 nCUBE-2 和 Intel iPSC 860, 分别可扩展到 8192 和 128 个结点, 峰值性能能达到 25GFLOPs 和 7GFLOPs。



图 1.3: Hardware-Controlled Thread Scheduling

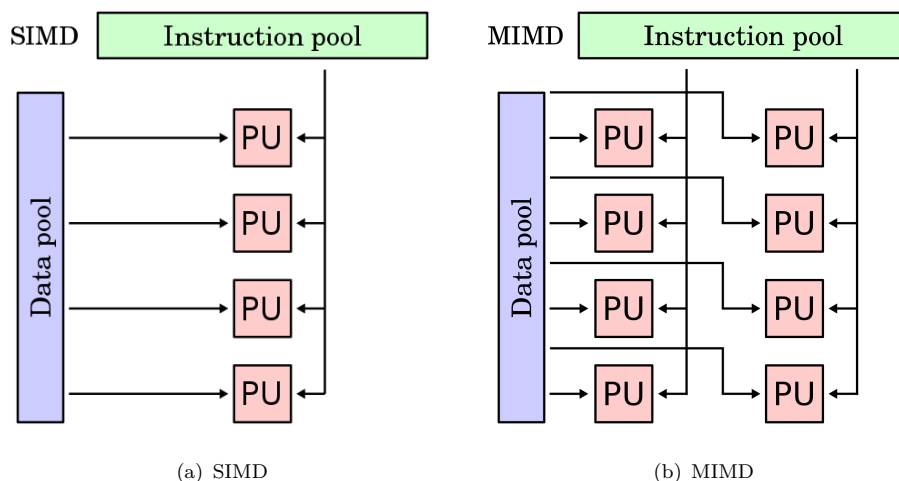
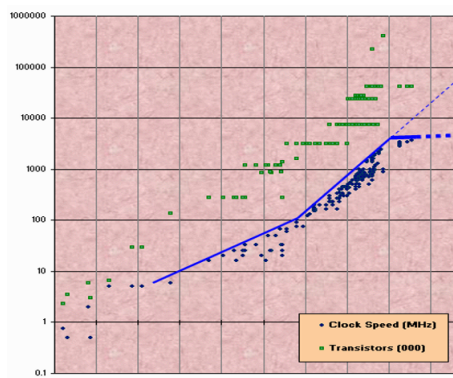


图 1.4: SIMD 和 MIMD 架构

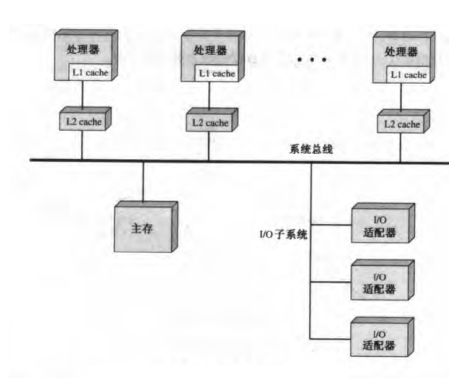
1.4 20 世纪 90 年代

90 年代得益于微电子技术的发展，基于 RISC 指令系统的微处理器芯片的性能几乎以每 18 个月增长 1 倍、内存容量几乎每 12 个月增长 1 倍的速度发展，而网络通信技术也得到了快速增长，它们都对并行计算机的发展产生了重要影响。而为了解决共享存储并行机不可避免的内存访问瓶颈问题，人们开始研究分布式存储 MMP（海量并行处理架构）。MMP 并行机大量采用商用微处理器芯片作为单结点，通过高性能网络连接而成。这一时期出现的分布式存储 MMP 并行机有：

- Intel Touchstone Delta，含 512 个 i860 微处理器，二维 Mesh 连接，峰值性能为 32GFLOPS，8GB 内存。
- CRAY T3D，16-1024 个结点，每个结点含有 2 个 64 位 RISC DECchip21064 处理器，峰值性能为 153GFLOPs。



(a) 集成度和主频



(b) SMP

90 年代中后期，微处理器的性能已经非常强大，能够提供每秒几亿到十几亿次的浮点运算速度。例如：

- IBM P2SC, 主频 135MHz, 峰值性能 500MFLOPs, “深蓝”超级计算机由此处理器构成。
- SUN Ultra SPARC, 主频 250MHz, 峰值性能 1GFLOPs。
- DEC Alpha 21164, 主频 600MHz, 峰值性能 1.2GFLOPs, 国产厂商“申威”的技术来源于此。

同时, 互连网络点对点通信能达到每秒超过 500MB 的带宽, 使得以高性能微处理器和互连网络通信为基础的共享存储对称多处理机 (SMP) 系统得到了发展, 并迅速取代了共享存储向量并行机的市场。此外, 由于分布式存储并行计算机具有并行程序设计困难、不易被用户接受等问题, 单一的分布式存储并行机已经朝分布共享 (DSM) 方向发展。

1.5 21 世纪至今

2000 年以来, 受重大挑战计算需求的牵引和微处理器以及商用高速互连网络持续发展的影响, 高性能并行计算机得到飞速发展。当前并行计算机的体系结构可以分为如下三类:

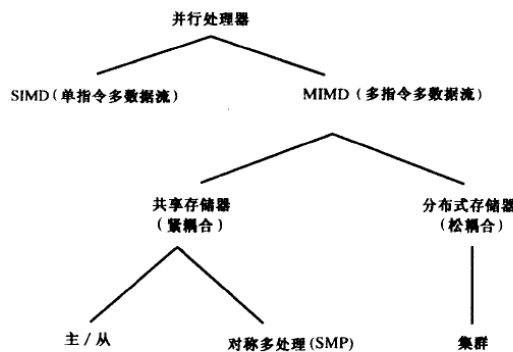


图 1.5: 体系结构分类

1. 集群 (Cluster)

集群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。这一类型的代表有:

- IBM JS20 集群 (2005 年 6 月 TOP500 第 5 名), 4800 颗 IBM PPC970 处理器, Myrinet 互连, 峰值性能为 42TFLOPs
- CDC Thunder (2005 年 6 月 TOP500 第 7 名), 4096 颗 Intel Itanium2 Tiger4 处理器, Quadrics 互连, 峰值性能为 22TFLOPs

2. 星群 (Constellation)

星群系统的每个结点是一个并行机子系统, 每个子系统里包含多个微处理器, 采用商用交换机连接, 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。这一类型的代表为 SGI Columbia 系统 (2005 年 TOP500 第 4 名), 20 个结点通过 Voltaire InfiniBand 网络连接, 每个结点为 SGI Altix3700 系统 (含由 SGI NUMalink 连接的 512 颗 Itanium2 处理器), 峰值性能为 60.9TFLOPs。

3. 大规模并行机系统 (MPP)

大规模并行机系统由结点构成, 每个结点含 10 个左右处理器, 共享存储, 处理器采用专用或商用 CPU, 采用专用高性能网络互连, 结点间分布存储。这一类型的代表有:

- IBM BLUE Gene/L (2005 年 6 月 TOP500 第 1 名), 65536 颗 IBM PowerPC440 双核处理器, IBM Proprietary 专用互连网, 峰值性能为 367TFLOPs
- NEC Earth Simulator (2005 年 6 月 TOP500 第 10 名), 5120 颗 NEC 处理器, 峰值性能为 40.9TFLOPs

2 中国超级计算机发展历史

2.1 银河一号

1978 年 3 月, 邓小平听取计算机发展汇报, 指出: “中国要搞四个现代化, 不能没有巨型机”, 明确由国防科工委系统承担首台亿次超级计算机的研制。同年 5 月, 中国超级计算机方案论证会上, 这项工程取名为“银河”。“银河”计算机规划性能为每秒 1 亿次浮点运算, 比此前中国研制的最先进的计算机还要快 100 倍。1983 年 12 月, 中国第一台亿次超级计算机“银河一号”通过国家技术鉴定, 横空出世, 这是中国超级计算机研制的一个重要里程碑, 标志着中国成为继美国、日本之后, 第三个能独立设计和制造超级计算机的国家。从此, 中国超级计算机的发展明显加快。

2.2 曙光系列

2004 年, 曙光 4000A 诞生, 2560 颗 AMD Opteron 处理器, Myrinet 互连, 峰值性能为 11TFLOPs, 在当年 6 月的 TOP500 中排名第十, 这是中国超级计算机首次进入世界前十。2008 年, 曙光 5000A 诞生, 6600 颗 AMD Opteron Quad 四核处理器, Infiniband DDR 互连, 峰值性能为 223TFLOPs, 在当年 11 月 TOP500 中排名第十一。曙光 5000A 的诞生, 标志着中国成为世界第二个有能力研制百万亿次超级计算机的国家。

2.3 天河系列

2009 年, 由国防科技大学和浪潮联合研制的天河一号诞生, 峰值运行速度 1203TFLOPs。2010 年, 第二期天河一号 A 研制成功, 其包括 7168 个计算结点和 1024 个服务结点, 内存总容量 262TB, 存储总容量 2PB, 每个计算结点包含 2 个 Intel CPU 和 1 个 NVIDIA GPU, 每个服务结点包含两个自研飞腾-1000CPU, 峰值性能 4701TFLOPs, 在 11 月的 TOP500 中排名第一, 这是中国自研的超级计算机首次夺得榜首, 同时天河一号率先采用了 CPU+GPU 异构融合的体系结构, 第一次从工程实现上证明了 GPU 可以被用于超级计算机。

2013 年, 由国防科技大学研制的天河二号诞生, 其由 16000 个运算结点组成, 每个结点配有 2 颗 Xeon E5 12 核处理器和 3 个 Xeon Phi 协处理器, 总计 312 万个计算核心, 总计有 1.404PB 内存和 12.4PB 外存。峰值性能为 54902TFLOPs, 在 2013 年 6 月到 2015 年连续 6 次在 TOP500 中排名第一。

2.4 神威·太湖之光

2016 年, 国家并行计算机工程技术研究中心研制的神威·太湖之光超级计算机诞生, 整个系统采用 40960 颗申威 SW-26010 260 核处理器, 总计 10649600 颗计算核心, 1.31PB 内存, 峰值性能 125PFLOPs, 比天河二号快了近 2 倍, 在当年 11 月以后的 4 次 TOP500 蝉联世界第一。

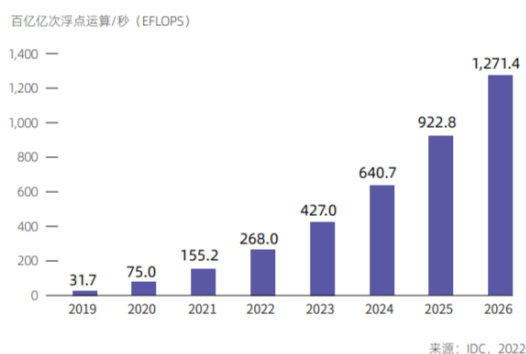
2.5 现状

中国将百亿亿次超级计算机的研究写入国家“十三五”计划中，中科曙光、国防科技大学和江南计算技术研究中心同时获批进行 E 级超算的原型机系统研制项目，形成中国 E 级超算“三头并进”的局面。中国在发布“神威·太湖之光”之后至今，并未向 TOP500 榜单提交测试更高性能的超级计算机。2018 年 7 月，“天河三号 E 级原型机系统”已经在国家超级计算天津中心完成研制部署，并顺利通过项目课题验收。天河三号原型机有 512 个迈创计算结点，128 个飞腾计算结点，浮点计算性能 3.146PFLOPs，并行存储总容量 1PB。2018 年 8 月，神威 E 级原型机系统在国家超级计算济南中心正式启用，该系统峰值性能 3.13PFLOPs。2018 年 10 月，由中科曙光牵头的曙光 E 级原型机系统完成交付。至此，“十三五”计划中 3 个不同技术路线的 E 级原型机系统正式交付。2023 年中国科学家投稿的项目入围戈登·贝尔奖，该项目是基于“神威·海洋之光”完成的。据该项目论文，海洋之光拥有超过 10 万个定制的申威 SW26010Pro 处理器结点，其理论峰值性能为 1.5EFLOPs。

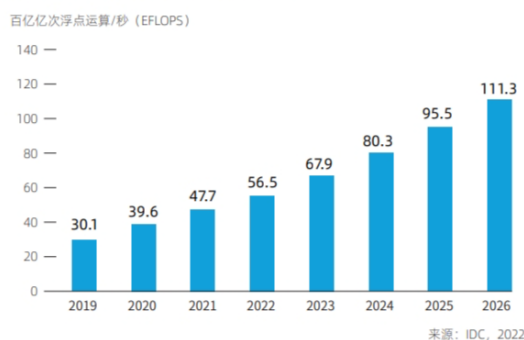
3 发展超级计算机的意义

2022 年，中国开始实施“东数西算”工程。“东数西算”是一盘关乎数字经济的大棋，相关机构测算的数据显示：区域算力指数平均每提高 1 点，其数字经济和 GDP 将分别增长 3.5‰ 和 1.8‰。超算作为“东数西算”的重要一环，将在其中发挥重要的算力支撑作用。当前，人工智能大模型上百亿、万亿规模的训练参数，对算力、数据、算法均提出更高的要求，而超算正越来越多地跟行业大模型深度融合，以算力的蓬勃发展激发经济增长新动能。算力是由市场驱动的，随着技术的进步，各个领域对算力的需求会越来越高。IDC 数据显示，2023 年中国智能算力规模达到 427EFlops，预计未来 5 年中国智能算力规模的年复合增长率将达 52.3%。大模型为各行各业带来海量 AI 算力需求，而大模型训练，将单一任务使用大量 GPU 卡运行数月，这是典型的超算应用特点，超算架构通过高速互联网络把 GPU 服务器紧耦合在一起，以并行计算技术实现单一大模型训练应用的高性能计算。因此，高性能、高质量、高性价比的超算架构成为赋能国产大模型、破解中国算力困局的刚需。

在今年在政府工作报告中，首次提到了算力，对算力有了明确的布局——超前建设数字基础设施，形成全国一体化算力体系。正如同工业社会离不开电一样，数字社会离不开算力，数字基建和算力是发展数字经济的命脉。因此大力发展超级计算机，对中国建设数字社会，算力强国至关重要。



(a) 中国智能算力规模及预测



(b) 中国通用算力规模及预测