



南开大学  
Nankai University

南 开 大 学

计 算 机 学 院

并行程序设计-体系结构调研报告

---

异构并行计算：从超级计算机到智能手机芯片

---

熊宇轩 2010056

年级：2020 级

专业：计算机科学与技术

指导教师：王刚

2022 年 3 月 5 日

## 摘要

异构并行计算，就是指将存在性能或架构差异的计算单元组成系统的计算方式。异构运算可以通过发挥系统中各种不同单元的特长，从而达到提升性能或者能效等目的。随着近年来对科学运算、AI 运算需求的提高，以及移动设备和物联网市场的扩大，异构并行计算愈来愈得到人们的关注。本文对近年来采用异构技术的超级计算机和芯片架构进行了简单剖析，阐述了异构并行运算是如何提升他们的性能或能效的，并且介绍了异构并行运算在软件设计方面带来的新挑战与解决办法。

**关键字：**异构超算；GPU；双调排序；ARM DynamIQ；英特尔 Alder Lake

## 目录

<b>一、 以性能为目标的异构并行计算机</b>	<b>1</b>
(一) 超级计算机的异构 . . . . .	1
(二) 异构在软件层面的挑战 . . . . .	1
(三) 并行排序算法——双调排序 . . . . .	1
(四) 矩阵乘法 . . . . .	2
<b>二、 以能效为目标的异构并行计算机</b>	<b>3</b>
(一) ARM . . . . .	3
(二) X86 . . . . .	3
<b>三、 总结</b>	<b>4</b>

## 一、以性能为目标的异构并行计算机

### (一) 超级计算机的异构

2008 年, 随着 680 块 Tesla T10 GPU 被安装到位于东京工业大学 Tsubame(意为“燕子”) 超级电脑, 超算榜单 TOP500 上首次出现了 GPU 的踪迹。GPU 虽然在通用运算能力上不及传统的 CPU, 但其具有的众核架构使得其在并行运算上比 CPU 快上一个数量级。GPU 这样的特性也 Tsubame 的算力升级到了 170 TFLOPS, 并成功跻身世界前三十强的超算。[1]

而随着时间的发展, 引入了 GPU 和协处理器的异构超算已经成为了高性能运算领域热门的选择。在 2019 年十月发布的 TOP500 榜单中, 排名前十的超级计算机, 有七台都是异构超算。在这七台异构超算中, 中国的天河 2 号使用了 Xeon Phi 31S1P 与 Matrix-2000 协处理器, NEC 的 SX-AURORA TSUBASA 则在芯片中集成了向量加速引擎, 剩下的五台使用则是英伟达 GPU。[7]

在相同功耗下, 异构超算在浮点性能上的收益非常之大。以天河 2 号和 2018 年到 2020 年位列 TOP500 榜单第一的 Summit(高峰) 超级计算机为例:

处理器	超算	类型	单节点	双精度浮点性能	功耗	每瓦双精浮点性能
			数量	/TFlops	/W	/GFlops/W
Xeon E5	天河 2 号	CPU	2	0.21	145	1.45
Xeon Phi		协处理器	3	1	300	3.33
Matrix-2000			3	2.4	200	12
POWER9	Summit	CPU	2	0.54	190	2.84
Tesla V100s		GPU	6	7.8	250	31.2

表 1: 超算处理器各项性能 [5] [8]

从表1可以看出, 无论是诸如 Xeon Phi 和 Matrix-2000 这样的协处理器, 还是 Tesla V100s 这样的 GPU, 其每瓦功能所能产生的双精度浮点性能都要比通用 CPU 高得多。

### (二) 异构在软件层面的挑战

虽然异构并行运算大大提升了超级计算机的性能和能效比, 但同时也不可避免地增加了软件设计方面的难度。以 CPU+GPU 异构的超算为例, CPU 所擅长运行的是串行的、对执行顺序有依赖的程序, 而 GPU 的优势则在于执行控制简单、高度并行的任务, 这就要求程序设计从传统的编程思维中跳出来。[8]

### (三) 并行排序算法——双调排序

以排序算法为例, 传统的排序算法诸如冒泡排序、快速排序为了使得数据有序, 需要运行大量的比较运算, 并根据比较运算的结果决定接下来的操作, 这样便在程序中产生了大量的分支和依赖, 使得排序算法很难在并行机器, 尤其是 GPU 这样众核架构的机器上达到高效率。当然, 也存在着许多并行的排序算法, 双调排序 (Bitonic Sort) 便是其中之一。

双调排序的核心是双调序列, 如果序列  $a_1, a_2, \dots, a_n$  是双调序列, 那么 (1) 序列中存在  $a_i$ , 使得  $a_1 \geq \dots \geq a_i \leq \dots \leq a_n$ ; 或者 (2) 序列循环位移后能够满足条件 (1)。

如果对长度为  $2n$  的双调序列  $X$  进行如下操作：(1) 将  $X$  分割成等长的  $MIN$   $MAX$ ；(2) 对所有  $MIN[i]$  和  $MAX[i]$  ( $0 \leq i \leq n-1$ ) 进行比较，如果  $MIN[i] \geq MAX[i]$ ，则交换两元素位置。操作后得到的序列  $MIN$  和  $MAX$  仍然是双调序列。这是双调序列的一个重要特性，被称为 Batcher 定理。如果对一个大的双调序列递归地进行这样的操作，直到  $n=1$ ，那么最终会得到一个有序序列。像这样将双调序列转换成有序序列的过程被称为双调排序 (Bitonic Sort)

而将双调排序这个过程的思路反过来，从单个元素开始，按照递增和递减顺序合并相邻的序列，就可以将无序的序列转换为双调序列，这个过程被称为双调合并 (Bitonic Merge)。[11]

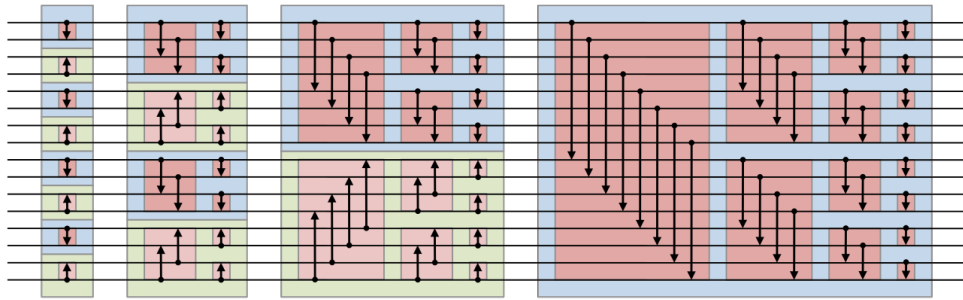


图 1: 双调排序的流程图 [2]

如图所示，双调排序是一个高度并行的算法，如果串行运行，其时间复杂度为  $O(n \log_2 n)$ ，而当有  $\frac{n}{2}$  个可用线程时，则其时间复杂度为  $O(\log_2 n)$ 。[2]

#### (四) 矩阵乘法

除了充分利用 GPU 之外，针对异构并行计算机编写的程序也需要考虑对各种架构进行充分地利用。在上文中提到的超算中，虽然 GPU 或协处理器提供了大部分的双精度浮点性能，但如果仅仅只将 CPU 作为 GPU 的控制器，所浪费的 CPU 性能也是非常多的。

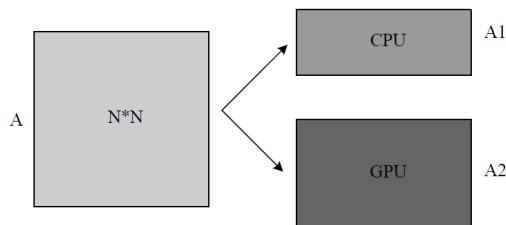


图 2: 通过预先的性能测量，确定矩阵分割的比例 [10]

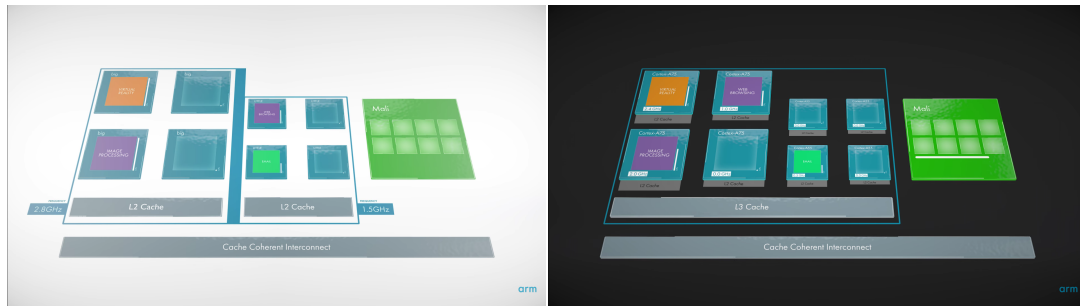
对于这种情况，可以通过性能分析，确定 CPU 和 GPU 求解问题效率的比值，并根据这个比值进行任务调度。比如，两个  $N \cdot N$  的矩阵  $A, B$  相乘，我们就可以将矩阵  $A$  按比例上下分割开来，形成矩阵  $A_1$  和  $A_2$ ， $A_1$  在 CPU 和  $B$  进行点乘运算， $A_2$  则和  $B$  一起发送给 GPU 进行点乘，这样既节省了数据在总线上的传输时间，也充分利用了 CPU 的性能。[10]

值得注意的是，最近几年，TOP500 登顶的两台计算机——神威太湖之光和富岳又回到了单芯片的路线，也就是说这些超算的节点中不再是 CPU+ 协处理器或 CPU+GPU 这样有多种处理器的配置了。但是这些超算仍然是异构的——神威太湖之光所使用的神威 26010 处理器在片上集成了融合异构众核架构、单芯片具有 4 个运算控制核心和 256 个运算核心；而富岳所使用的 A64FX 也是 48 个计算核心、4 个辅助核心的异构结构。总而言之，单芯片超算是对异构并行计算架构的一种肯定与发展。[9] [4]

## 二、以能效为目标的异构并行计算机

### (一) ARM

作为在便携设备上得到广泛应用的架构，ARM 架构的功耗控制极为优秀，而近年来智能手机市场的蓬勃发展，对 arm 架构的性能和效率提出了更高的要求。为了满足这样的要求，ARM 公司于分别于 2011 年和 2017 年推出了异构多核的结构配置方案——big.LITTLE 和其升级版 DynamIQ。



(a) ARM big.LITTLE 架构

(b) ARM DynamIQ 架构

图 3: ARM 公司推出的两种异构多核配置 [3]

DynamIQ 相较于 big.LITTLE，最大的优点就是新增了可以更自由配置的大小核心数量的特性。在 big.LITTLE 中，相同的核心被打包在同一个集群（Cluster）里，一种典型应用是，四个大核（Cortex-A15）和四个小核（Cortex-A7）被分别打包为两个集群（Cluster），而两个集群之间则通过一致性互联总线（Coherent Interconnect）链接。而到了 DynamIQ 中，单个集群中核心的数量变为了 8 个，且不同的核心可以被打包在一个集群里。这样一来，ARM 的客户便可以根据自己的需求自行配置大小核的比例 [3]

大小核的架构也对软件的设计提出了更高的要求。操作系统需要了解当前的电量状态，每个线程的性能需求等信息，并根据这些信息将线程转移到适合它们的核心上去，只有合理的调度算法，才能发挥大小核架构在效率方面的潜能。

### (二) X86

随着近年来笔记本电脑市场的扩张，在 PC 市场具有支配性地位的 X86 处理器设计公司们也愈来愈需要关注其产品能耗比这项指标了。英特尔公司在其最新发布的 Alder Lake 架构中，也使用了异构多核的技术。Alder Lake 具有两种核心——被称为 Golden Cove 的性能核心，以及被称为 Gracemont 的效率核心。

英特尔 Alder Lake 的亮点在于其调度技术——Thread Director。当一个任务需要较少的性能时，会被优先分配到效率核心上，反之，则会被分配到性能核心上。而当所有的性能核心都被占用时，出现了一个需要更高性能的任务，比如说需要性能核心向量处理性能的 AI 任务，Thread Director 便会向操作系统建议，将一个在性能核心上运行的任务转移到效率核心上。此外，如果性能核心上出现了等待状态的线程，它也会被移动到效率核心上运行。[6]

### 三、 总结

异构计算机在提升超级计算机性能、帮助移动设备获得更长续航时间这两件事上已经表现出了其巨大的潜力。我们应该相信，随着相关软硬件技术的发展，异构计算机将会变得更快、更强、更环保。

NIKU

## 参考文献

- [1] Tsubame(supercomputer). [https://en.wikipedia.org/wiki/Tsubame\\_\(supercomputer\)](https://en.wikipedia.org/wiki/Tsubame_(supercomputer)), Dec 2020.
- [2] Bitonic sorter. [https://en.wikipedia.org/wiki/Bitonic\\_sorter](https://en.wikipedia.org/wiki/Bitonic_sorter), Dec 2021.
- [3] ARM. Arm dynamiq redefines multi-core computing. [https://www.youtube.com/watch?v=qPGTP\\_ZxDyY](https://www.youtube.com/watch?v=qPGTP_ZxDyY), Sep 2017.
- [4] FUJITSU. A64fx. <https://www.fujitsu.com/global/products/computing/servers/supercomputer/a64fx/>.
- [5] Oak Ridge National Laboratory. Summit. <https://www.olcf.ornl.gov/summit/>.
- [6] Intel Technology. Alder lake and thread director - architecture day 2021 | intel technology. <https://www.youtube.com/watch?v=lm6vovVFo70>, Aug 2021.
- [7] Weimin ZHENG. Research trend of large-scale supercomputers and applications from the top500 and gordon bell prize. *Science China(Information Sciences)*, v.63(07):128–141, 2020.
- [8] 张竞扬. 为何中国超算偏爱异构计算. <https://zhuanlan.zhihu.com/p/20908218>, May 2016.
- [9] 成都申威科技有限责任公司. 申威 26010. <http://www.swcpu.cn/show-130-261-1.html>.
- [10] 方旭东. 面向大规模科学计算的 CPU-GPU 异构并行技术研究. PhD thesis, 国防科学技术大学.
- [11] 谢小帅. 双调排序. <https://www.jianshu.com/p/ea4a62fdaae9>, Sep 2017.