

高性能计算的发展

臧大伟, 曹政, 孙凝晖

中国科学院计算技术研究所, 北京 100190

摘要 继理论科学和实验科学之后, 高性能计算成为人类科学研究的第三大范式。作为科技创新的重要手段, 高性能计算广泛应用于核爆模拟、天气预报、工程计算等众多领域, 是当代科技竞争的战略制高点, 集中体现一个国家的综合实力。本文介绍高性能计算发展的历史和现状, 分析当前高性能计算所面临的问题和挑战, 探讨高性能计算未来的发展方向。

关键词 高性能计算; 高性能计算机; 高性能计算市场

利用大量处理单元的聚合计算能力来解决复杂问题, 是高性能计算 (high performance computing, HPC) 最直观的定义。高性能计算已成为继理论科学和实验科学之后科学探索的第三范式, 被广泛应用在高能物理研究、核武器设计、航天航空飞行器设计、国民经济的预测和决策、能源勘探、中长期天气预报、卫星图像处理、情报分析、互联网服务、工业仿真等领域, 对国民经济发展和国防建设具有重要的价值。它作为世界高技术领域的战略制高点, 已经成为科技进步的重要标志之一, 同时也是一个国家科技综合实力的集中体现。

本文介绍高性能计算的发展现状, 分析现在高性能计算面临的挑战, 探讨中国高性能计算未来的发展方向。

1 高性能计算的发展现状

高性能计算作为计算机科学的一个分支, 致力于开发高性能计算机和运行在高性能计算机上的应用软件。回顾历史, 高性能计算作为一个强大的计算工具, 与科学研究的发展密不可分。一方面, 科学研究对计算能力永无止境的需求促进了高性能计算技术向前发展; 另一方面, 高性能计算技术的每一次巨大进步都为科学研究提供了全新

的手段。

1.1 永无止境的计算需求

在近代科学研究中, 单靠理论和实验解决问题的难度逐渐增大, 数值运算的方法被用来模拟物理世界, 以求解复杂的问题, 计算科学成为自然科学研究的必备工具。随着求解问题规模的越来越大, 对计算能力的需求成为驱动高性能计算发展最直接的动力。

第二次世界大战时期, 靠人力计算火炮的弹道非常困难, 战争对计算能力的需要促进了第一台电子计算机的诞生; 早期的高性能计算机主要应用于解决军事领域的计算问题, 如美国在 1960 年代使用 CDC 超级计算机进行弹道计算、火箭设计等工作^[1]。20 世纪 90

年代中期以后, 随着机群技术构建的高性能计算机的普及, 高性能计算的成本和编程的复杂度大幅度下降, 为高性能计算的广泛使用创造了条件。如图 1^[2]所示, 现在高性能计算已经渗透到各个学科领域, 不仅在气候模拟、石油勘探、天体物理这些传统应用领域保持强劲的生命力, 在生命科学、人工智能、大数据处理这些新兴领域也有广泛的应用。

在传统应用领域, 如天气预报、石油勘探、核爆模拟等, 计算问题一般采用划分网格的方式来解决, 随着应用的物理建模不断精细, 数值模拟分辨率越来越高, 对计算能力的要求也越来越高。以气候模拟与天气预报为例, 建立模拟气候变化的模型是一个非常挑

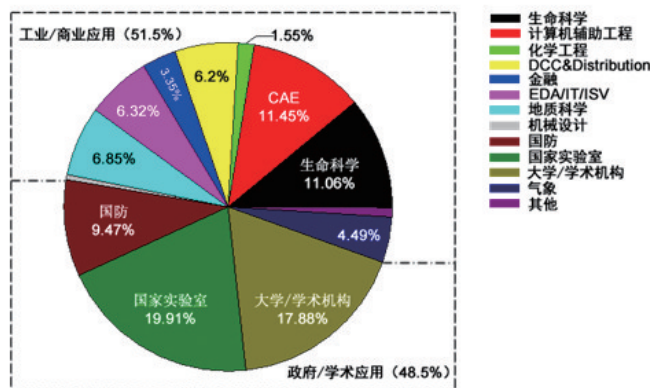


图1 2014年全球高性能应用领域分布

收稿日期: 2016-05-30; 修回日期: 2016-07-07

作者简介: 臧大伟, 助理研究员, 研究方向为高性能计算、数据中心网络, 电子邮箱: zangdawei@ncic.ac.cn

引用格式: 臧大伟, 曹政, 孙凝晖. 高性能计算的发展[J]. 科技导报, 2016, 34(14): 22-28; doi: 10.3981/j.issn.1000-7857.2016.14.002

战性的工作,它需要模拟巨量的实体间相互作用,同时还需要在不同的时间和空间维度上进行分析,一般使用数十亿个非线性方程从不同的维度描述各种物理过程。如图2^[3]所示,现在比较成熟的全球气候模拟模型一般采用100~200 km的网格,它对地形效应、细粒度水文状况的模拟能力很弱;最新出现的模型采用20~50 km的网格,在此分辨率下可以得到很大的改善,但需要强大得多的计算能力才能求解。如果将每个空间维度上的分辨率再提升1倍,则总的计算能力至少需要提升8倍。据分析,如果将模型的分辨率提升到1 km,则相应的计算能力需要提升100~1000倍,这需要未来将高性能计算机的计算能力由现在的P级升级为E级(10^{18})。

天气预报一般采用更高级的对流解析有限域模型,在未来很长一段时间内,实现1 km水平分辨率的对流运动模型是气象领域的一项重要工作。高分辨率的模型可以直接地求解对流系统的问题,模拟结果可以更好的展示地形效应、海洋大气能力转移过程,并且可以获得更详细的区域气候情况。这些高分辨率的模拟结果能帮助我们更好地理解全球变暖对天气的影响,同时可以利用对区域气候的模拟来评估极端天气事件对社会的影响。

激光聚变数值模拟对计算能力有着相似的需求。为了模拟内爆过程中辐射流体力学界面不稳定性演化,100个波长的模拟是最低的要求。对单模模型,每个波长至少需要10个网格,则单个方向需要1000个网格,三维模拟总共需要 10^9 个网格。对多模模型,为了分辨单个扰动模,每个波长需要50个网格,如果模拟200个波长,需要的网格总数约为 10^{12} 个^[4]。当前,千万亿次计算机的CPU核数为 $10^4 \sim 10^5$ 量级,勉强可满足单模模型的需求,但模拟多模模型是目前千万亿次计算机难以承受的,计算能力至少需要有两个量级的提升,需要更高性能的计算机。

在非传统的新兴应用领域,如生命科学、人工智能、大数据处理,这些应用

的负载很多都是基于图模型和图算法来处理数据,而复杂的图结构的规模非常庞大。例如,在娃娃鱼基因组测序中,对应De Bruijn图有超过 10^{13} 个顶点,测序技术的发展对计算能力的需求不断提高。

深度学习是新兴领域中另一个典型代表。深度学习技术试图通过大规模的神经网络和大数据提供的海量训练集合,将大脑学习识别的过程加以抽象,从而获得极高的识别准确度,这些都带来了极大的计算需求和吞吐需求。在早期使用深度神经网络进行语音识别的模型中,拥有429个神经元的输入层,整个网络拥有156 M个参数,训练时间超过75天^[5,6];人工智能和机器学习顶级学者Andrew Ng和分布式系统顶级专家Jeff Dean打造的Google

Brain项目^[7],用了包含16000个CPU核的并行计算平台训练超过10亿个神经元的深度神经网络,在语音识别和图像识别等领域取得了突破性的进展。如图3^[8]所示,如果模拟人类大脑的全脑级的神经系统,需要模拟1000亿个神经元,需要计算能力有数个数量级的提升。许多商用或开源系统如Caffe、Theano、TensorFlow等相继出现,尤其是由谷歌开源的TensorFlow系统,支持异构设备分布式计算,可以灵活的从单个CPU/GPU扩展到数千个CPU/GPU组成的分布式系统;深度学习的进步也促进了专用加速器的发展,例如中国科学院计算技术研究所陈云团队所设计的寒武纪神经网络处理器,相对于x86指令集的CPU有两个数量级的提升,而面积和功耗只有其1/10。

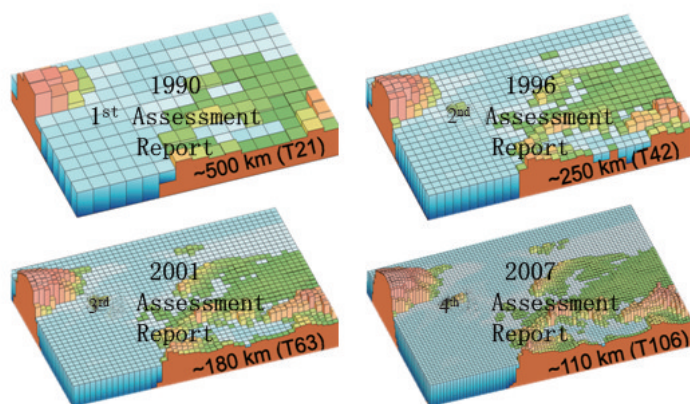


图2 IPCC评估报告中所使用气候模型的空间分辨率

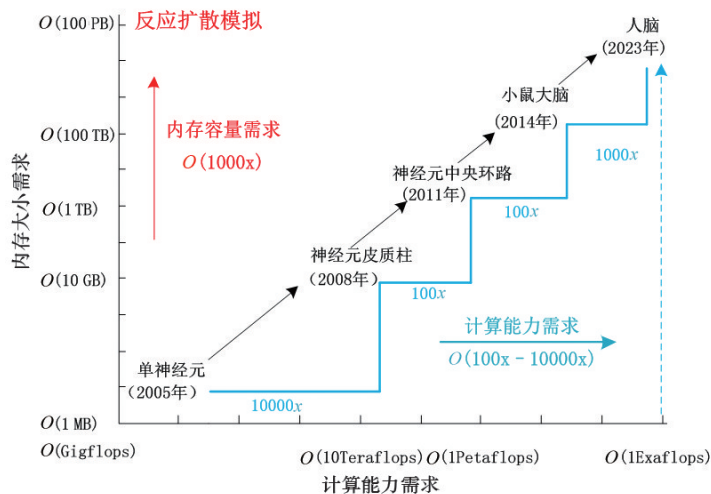


图3 脑模拟对高性能计算性能的要求

以上所列举的各个学科对计算能力的需求有一些相似的特点:它们需要强大的计算能力来模拟更大规模的应用,并同时增加应用的分辨率,因此需要的计算能力要有几个数量级的提升。按照推测,满足这些应用的未来计算机系统峰值性能在2020年至少应该达到1 Eflop/s。

1.2 计算利器:高性能计算系统

高性能计算系统利用大量处理单元的聚合计算能力来满足应用巨大的计算需求,其关键问题是实现众多计算节点的大规模集成和高效协同计算,核心技术涉及高性能计算机和大规模并行应用程序。

1.2.1 高性能计算机

自第一台电子计算机ENIAC问世以来,以电子器件、系统结构和计算模式的重大变革为标志,高性能计算机发展已经历经三次大的发展阶段(图4)。

集成电路的发明推动了信息技术革命,1965年,Intel创始人Moore提出集成电路的第一定律“摩尔定律”,即“集成电路上可容纳的晶体管数目大约每隔18个月便会增加1倍,性能也提升1倍”^[9]。摩尔定律完美预测了此后近50年通用微处理器的发展轨迹,但“摩尔定律”的提升速度不足以满足快速增长的计算需求,体系结构开始引领高性能计算机的发展,相继出现了向量机、SMP、ccNUMA、MPP(massively parallel processing)和机群5种主要架构。

“扩展性”和“成本”是推动体系结

构进步的两个关键因素。从20世纪70年代中期到90年代初的接近20年时间里,向量机占据了高性能计算机的统治地位。虽然仅一条指令就可以处理整个向量,但向量化编程存在困难,加之全定制处理器高昂的成本,导致向量机遇到了瓶颈,基于通用微处理器构建的共享存储多处理器系统(SMP)逐渐成为主流。SMP集中式的共享机制导致了扩展性受到限制,带有分布式特征的ccNUMA(支持cache一致性的非一致访存架构)被斯坦福大学提出,并在1991年完成了第一台ccNUMA架构的并行机Dash。ccNUMA结构延续了SMP的编程模型,苛刻的远程访存延迟需求决定了其难以实现大规模的扩展。

为了解决大规模扩展的问题,很快出现了MPP体系结构,它可扩展到上万个结点。其结点被分为计算结点和 service 结点两类,计算结点运行轻量级定制操作系统,用于提供计算能力; service 结点运行完整操作系统,为计算结点提供诸如文件系统、任务管理和I/O等服务。结点间以定制的高速网络互连,采用基于消息传递的编程模型。因为MPP采用了专用部件,较高的构建和维护成本成为它广泛使用的限制因素。Cluster(机群)应运而生,其节点和网络均采用商业化的部件,制造和维护成本都很低,同时具备MPP的高可扩展特性,迅速成为高性能计算机领域的主流。到2007年,世界Top 500榜单的系统就已经完全被MPP和Cluster系统垄

断(图5)。

但是到了2008年左右,能耗问题制约了处理器主频的大幅度提升,如果继续采用当时主流的机群技术(曙光4000A, Top 500的第10名),实现千万亿次系统需要64000个节点,占地近14000 m²(约两个足球场),功耗约38 MW(一个中等县城的用电量)。为了突破系统规模、计算密度、系统能耗这三个因素的制约,采用通用处理器和加速器协同计算的异构计算模式被提出,并成为系统性能提升的主要推动力,其中主流的加速器有GPGPU、众核处理器和FPGA。由于加速器不需要执行复杂的控制流,专注于浮点计算,因此它可以实现大量浮点计算部件的堆叠,峰值计算性能超过通用处理器数倍或数十倍,对一些特定应用甚至可以实现上百倍性能的提升,极大提高了计算节点的效能比。在2015年11月发布的Top 500排名中^[10],使用加速器的系统提供的总性能达到榜单中计算机性能总和的1/3,新构建的高性能计算机系统中,近1/2采用加速器,目前已有超过2/3的高端应用代码支持GPU优化。

图6是支持异构计算的曙光高性能计算机结构示意图,其高密度异构计算刀片节点可以继承大量高性能异构计算部件,节点间采用大规模互连网络连接,并配以海量存储系统。中国的曙光6000(星云)和天河1号系统分别使用了NVIDIA和ATI公司的GPU作为加速器,在2010年先后获得了世界高性能

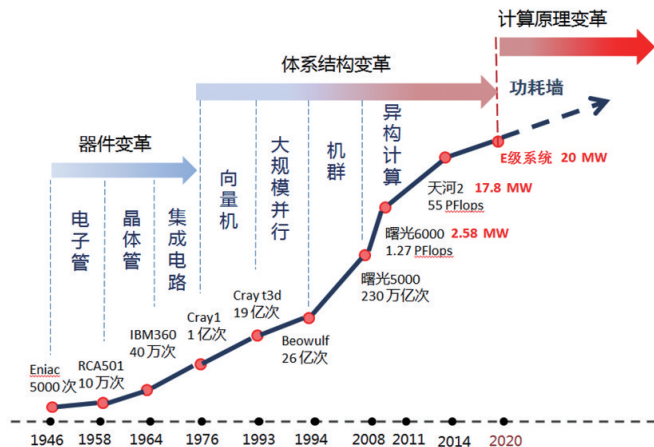


图4 高性能计算机发展历史

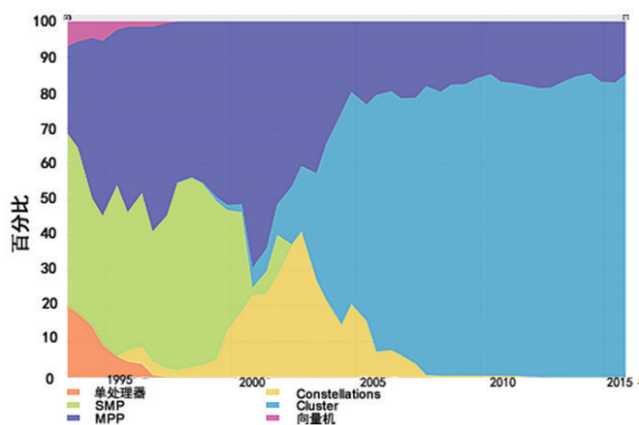


图5 高性能计算机体系结构发展

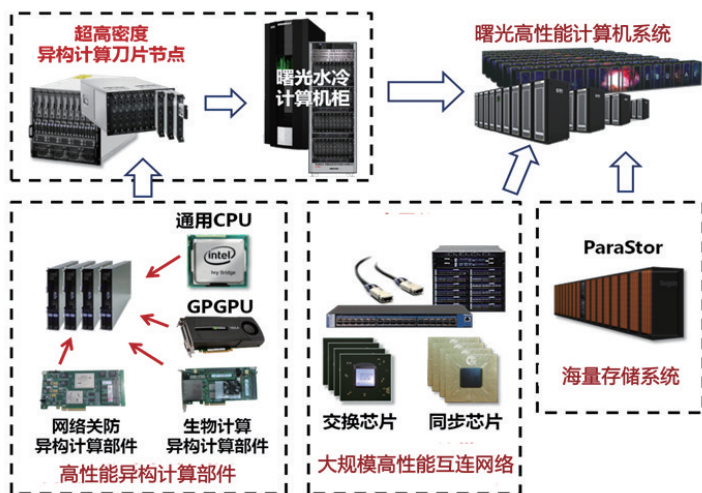


图6 曙光系列高性能计算机组成结构示意图

计算机 Top 500 榜单的第二和第一名, 采用 Intel Xeon Phi 众核加速器的天河 2 号系统自 2013 年就一直保持着世界第一的排名。

2015 年美国总统奥巴马签署推出“国家战略计算计划”(NSCI)的行政命令来支持高性能计算机的继续发展, 目标建造计算峰值达到百亿亿次的高性能计算机, 也就是 E 级计算机系统^[11]。中国也启动了在“十三五”期间建造 E 级系统的计划。异构计算模式将是达到 E 级计算能力的必由之路, 这一点国际上已经达成共识; 由于系统功耗限制(功耗墙), E 级系统之后, 仅借助当前的异构计算模式难以再实现大幅性能提升(图 4)。人们寄希望于计算原理的突破来开启高性能计算机的下一个大发展阶段, 量子计算、光计算和 DNA 计算都是这一方向的研究热点。2013 年美国谷歌、NASA 和大学空间研究联合会购买了 D-Wave 公司的“D-Wave Two”量子计算机, 成为量子计算产业化的标志性事件。但量子计算机仍处于探索阶段, 存在大量结构和算法问题, 短期内难以规模产业化。

1.2.2 基础算法库和并行应用软件

高性能计算机的峰值性能不等于应用软件运行时的实际性能, 它们之间往往存在巨大的鸿沟。以稀疏矩阵类应用为例, SpMV(稀疏矩阵乘法)的浮点性能通常不超过峰值性能的 10%,

在 GPU 众核处理器上的浮点效率甚至低于 2%。要发挥高性能计算机的高速硬件优势, 必须要有适用的算法和调优的应用程序来实现数百万核之间的并行。因此, 大规模并行软件和高性能算法的发展水平象征着各个国家高性能计算的软实力。

现在大规模并行软件在各个领域发挥着重要作用, 知名的有: 大气领域的 WRF、计算化学的 Gaussian、流体力学的 Fluent、LS-Dyna 等。美国和日本是高性能并行软件强国, 高性能计算领域最高奖项“戈登·贝尔奖”20 多年来一直被美国和日本垄断; 欧洲同样非常注重应用软件的研发, 计算化学的 ADF、MOLPRO、分子模拟的 GROMACS、材料计算的 VASP 在世界范围广泛使用。大规模并行应用软件的发展都与国家巨额投入息息相关, 美国能源部(DOE)在硬件上的花费不到总投资的 1/6, 大部分预算都花在了物理建模、算法研究和软件研制方面。

国内则更重视有显示度的高性能计算机硬件的研制, 对应用软件的投入欠缺, 且缺乏整体计划, 虽然也开发了数个几十万核、乃至百万核的大规模应用, 但多局限于对于计算数据的测试、算法程序的并行优化等基本的辅助性操作。

大规模并行应用程序的核心是基础算法模块, 许多科学问题的解决高度

依赖于基础算法与可计算建模的发展水平。高效的基础算法和满足实际精度要求的可计算模型可以显著降低计算复杂度和计算量, 提高利用计算机解决科学与工程问题的能力。例如, 著名的高性能数学库 BLAS、ScaLAPACK 和 FFTW 等在提高大量应用性能上发挥了关键性作用。2012 年 3 月美国能源部发布了题为“Report on the extreme-scale solvers: Transition to future architectures”的报告, 指出在 CPU 核数为 10 万量级的计算机上, 稀疏线性解法器在很多复杂应用数值模拟中占了 90% 的时间; 在核爆模拟和激光聚变等很多数值模拟应用中, 稀疏线性解法器也同样消耗了绝大部分运行时间。美国能源部“先进计算促进科学发现(SciDAC)”项目在最新发布的第 3 期计划中, 成立了 FASTMath (Frameworks, Algorithms, and Scalable Technologies for Mathematics) 研究小组作为第一批启动的重点内容, 其目标是面向实际复杂应用的大规模数值模拟, 发展可扩展的共性算法和使能技术, 最终形成高性能数学工具箱, 包括了 13 个具有共性的软件包。

北京应用物理与计算数学研究所和中国科学院数学与系统科学研究院分别研制了 JASMIN 框架和 PHG 平台, 面向科学计算领域中的自适应结构网格和非结构网格数值模拟应用, 它们将科学计算中现有的很多共性算法集成, 并封装形成共性层模块, 支撑数值模拟应用(如激光聚变、油藏和电磁场)在国产高性能计算机上的发展。

在互联网领域, 大数据处理应用也存在共性基础模块。例如, 数据挖掘中大量采用的线性代数解法器、网页排序算法 PageRank 属于典型的稀疏线性迭代方法。在社交网络分析、系统生物学和基因测序中, 基于图的建模是基本的处理方式, 图算法已经成为事实上的基础模块。这些实际问题中的图多数用稀疏矩阵来描述, 数学模型可抽象为线性代数的表示, 其算法与稀疏线性代数存在数学上的等价性。这类数据处理应用具有更严重的不规则计算与通信

模式,以及更低的计算访存(通信)比,如何并行优化成为高性能计算研究的热点和难点。2010年6月,美国Sandia实验室牵头联合美国几大国家实验室和科研机构,发布了以图遍历算法为核心的Graph500基准测试程序^[12],在用于高性能计算机系统排名的同时,也促进了面向E级数据处理的并行算法设计和优化技术的研究。

1.3 高性能计算机产业

根据IDC 2015年的统计^[12],全球高性能计算市场规模在250亿美元,其中高性能计算机系统(包括服务器、存储和网络)约占60%,软件和服务约占35%;据预测在2015—2020年间高性能计算市场规模将以8.3%的复合增长率而增长,在2020年达到440亿美元。自2012年的4年间,世界高性能计算机Top 500排行榜的入门性能和性能总和分别提高了4倍和3.7倍。

美国公司仍然占据整机市场的领先地位,2014年全球Top 500高性能计算机市场份额的80%被惠普、IBM和Cray三家公司占据(图7)。但受系统升级以及2014年联想收购IBM的x86服务器部门的影响,2015年IBM的份额大幅下降,国产的曙光高性能计算机异军突起,以9.8%的份额取代IBM位列第三,这证明了中国高性能计算机整机技术和产业化能力都达到了世界领先水平。在国家高技术研究发展计划(863计划)的持续支持下,中国已经掌握了包括高密度服务器、机群操作系统、高性能存储系统、冷却技术在内的各项整机系统技术。国产高性能计算机系统已经连续两年占据中国Top 100系统90%以上的份额,曙光信息产业股份有限公司更是连续7年蝉联中国Top 100份额第一。随着联想对IBM x86服务器并购的完成,曙光、联想和浪潮三强争霸的局面已经形成。

在国家自主可控和保障信息安全的宏观政策引导下,中国启动了基于自主可控处理器构建国产高性能计算机的计划,“核高基”重大科技专项支持了3个高性能处理器系列的研制,即飞腾(国防科大,ARM指令集)、申威(江南

计算所,基于Alpha的自定义指令集)和龙芯(中科龙芯,MIPS指令集)。国家高技术研究发展计划(863计划)项目中1Flops神威蓝光计算机全部采用了申威16核CPU,曙光星云计算机部分部署了龙芯8核CPU,国防科技大学天河系统部分部署了飞腾多核CPU。但由于采用非x86指令集,商用并行软件都不能运行,还没有能力构建自己的软件生态系统,这些因素使其应用领域受到很大限制,市场容量过于狭小。到目前为止,基于自主可控处理器的高性能计算机还局限在国家科技项目范畴,尚没有开始产业化进程,在国家战略应用中进行推广是一条可行的发展道路。

国产HPC应用软件是中国高性能计算的短板,与国际上的差距最大。在工业界应用极为广泛的CAE软件,仍然被以ANSYS等为代表的国外软件垄断,几百并行度的软件就需要数百万人民币的License费用,高端的开源HPC软件也主要来自美国、日本以及欧洲国家。具有代表性的国产商用应用软件仅有GeoEast(石油勘探领域)和Grape(大气科学领域),由于投入不足,近年来没有再涌现类似商用软件。大多数的国产私有应用软件也是以大学和科研院所自用为主,少有成规模的推广使用,特别是国产HPC应用软件在开源和中间件上基本刚起步,而它们早已成为应用软件规模化发展的主流模式。

2 高性能计算面临的挑战

根据Top 500历史数据进行预测,高性能计算机将在2020年左右进入E级时代。美国政府在“Strategy for

American Innovation”计划中,将E级计算列为21世纪美国最主要的技术挑战,美国国防部、能源部、自然科学基金委、国家核安全局等部门以及日本文部省、欧盟、俄罗斯联邦原子能署等均对此进行了大量的投入,中国也在“十三五”期间启动了E级计算机的研制计划。E级计算系统在能耗、可靠性、应用效率等方面都将面临一系列挑战。

2.1 功耗挑战

美国、日本、欧盟已经制定了2018—2020年研制E级高性能计算机的目标,系统功耗指标设定为20 MW,即每瓦500亿次浮点计算,而目前能效比最高的系统,即Green 500排名第一的由日本研制的Shoubu超级计算机,其能效比也仅达到70亿次/w的水平^[13],距离E级机的能效指标尚有10倍左右的差距。可以说,制约E级机实现的最大技术障碍之一是能耗比。

针对降低高性能计算系统能耗这一关键问题,国际学术界和工业界已有大量努力,涉及计算机系统从应用至芯片各个层次的优化,例如高效能众核处理器、减少数据移动的算法优化、能耗感知的系统调度、低能耗的系统、与机房冷却技术等。从目前的技术水平估计,即使考虑“摩尔定律”因素,在2020年实现20 MW的功耗目标仍存在相当大的困难。

2.2 可靠性的挑战

随着高性能计算机规模越来越大,软件结构越来越复杂,E级系统中的故障检测与诊断是非常有挑战性的问题。在亿亿次的规模下,系统的平均无故障时间(mean time between failure,

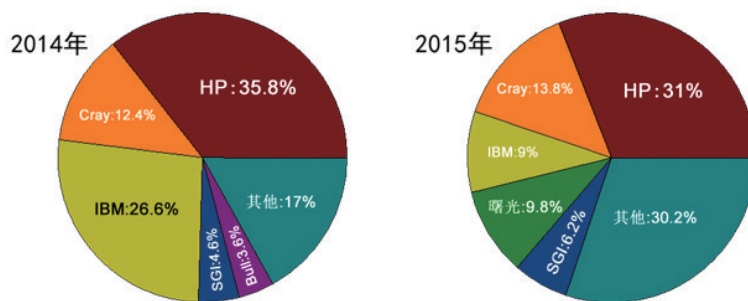


图7 全球Top 500高性能计算机市场份额

MTBF)仅为5 h左右;而在E级系统中,处理器的核数将达到 10^8 的量级,在概率上MTBF将会变得更短,其可靠性问题将会更加严重。越来越短的平均无故障时间导致故障将在海量现场数据的保存或恢复过程发生,传统的周期性保存现场的粗粒度检查点机制将会失效,在算法和应用层次进行细粒度容错设计,将是解决可靠性问题的重要途径。

由于涉及海量信息(包括各种日志、系统运行性能指标)采集和多维度(硬件、软件、误操作等)故障分析,及时定位故障也变得越来越困难。需要提高故障检测与诊断的能力,使得E级计算机系统能够快速发现故障,避免故障在系统中扩散,诊断出造成故障的根源,从而加速系统的恢复,保障系统的可用性。

2.3 应用效率挑战

为降低系统的能耗,构建E级计算机系统将大量使用GPU或众核处理器。目前,GPU或众核处理器都已经拥有数百计算单元,E级计算系统总体并发度将达到数千万量级,异构特性和海量并发度给并行程序的编写、调试、性能调优带来了巨大的挑战,极大地加剧了峰值性能与应用性能之间的鸿沟,E级计算系统可能仅能发挥出1%的峰值性能,应用效率变得极低。

由于极大规模并行所带来的复杂性,编程模型必须能够表示所有异构层次的内在并行性和局部性,以实现可扩展性和可移植性。同时,编程范式要能充分利用分布存储机制,以减少数据移动的开销。面向领域的编程框架和算法工具箱有可能成为提高产出率的有效途径。

目前中国在E级计算机研究方面与美国相比还存在较大差距,主要体现在基础技术储备不足,核心关键技术难以满足E级计算的需求。应力争在主流技术路线的若干关键技术和重大应用上取得突破,把握未来5~10年的关键时期,保持中国在高性能计算技术方面的可持续发展,更好地支撑中国科学研究走向世界一流,并进一步提升中国

高性能计算机产业的国际竞争力。

3 高性能计算的未来方向

3.1 以应用为导向发展高性能计算机

当前的通用系统架构以一种结构应对多种计算需求,越来越无法实现应用需求与硬件性能的最优适配,导致计算的低效和功耗的浪费。协同设计(co-design)正在被越来越多的专家认为解决这一问题的主要技术途径和顶层方法论。所谓协同设计是指通过领域科学与计算机科学间的跨学科紧密协作,抽象出面向领域应用的负载特征,根据应用的特征对计算、访存、通信能力进行优化配置,实现自硬件到应用软件的一体化定制设计。

应用与系统的协同设计理念使得E级系统的设计更加具有针对性,E级计算应用协同设计方法,需要覆盖气候模拟、核聚变、天体物理、材料科学、生物信息和人工智能等计算相关学科,从应用物理模型、计算方法、并行软件实现等多个层面寻求应用共性、区分特性,将系统研制与多领域应用软件能力提升拧成一体,最大限度发挥未来计算系统的通用计算效能。随着系统效能问题的日益严峻,领域定制系统将带动面向典型负载的新型加速器和处理器的发展,面向领域定制的高性能计算机有望成为未来高端HPC市场的主流。此外,领域定制系统将带动面向典型负载的新型处理器市场,成为国产处理器获得发展的机会。另外,中国的并行应用软件起步较晚,遗产代码量相对较少,反而成为中国发展自硬件到应用的全定制HPC系统的优势。

3.2 面向新兴应用发展新型高性能计算机

随着互联网的普及和技术的发展,许多与传统高性能计算应用完全不同的应用模式竞相出现,如Web服务应用、物联网服务应用、云计算应用等,在Google、百度、阿里等互联网企业使用数十万的服务器向数以亿计的用户提供各种服务。这些新型应用很多是基于海量数据提供吞吐密集型服务,这与传统计算密集型的高性能应用有很大

的差别,这从另一个方面赋予了高性能计算新的含义,即面向服务的高通量计算(high-volume throughput computing, HTC)。

高通量计算一般是基于海量的数据向大量的用户提供交互式、高并发的服务,在用户和负载动态变化时能够动态的扩展以满足对系统处理能力的需要,同时这样的计算对成本非常的敏感。但是现在运行这些应用的计算机系统体系结构本质上与用于科学和工程计算的系统是相同的,完全是通用处理器加通用系统技术,成本高、效能低。例如,现在面向Web服务的应用计算量很少,对处理器的浮点部件要求较低,而对I/O系统的要求较高,如果使用通用部件会造成成本的极大浪费。

现在国际上对于这种面向特定领域的高通量计算的研究还处于起步状态,对于技术路线和标准的制定还处于碰撞期。随着互联网的进一步发展,这必然会发展成为另一个战略高地,因此中国应该及时加大这个方向的投资力度,解决高通量专用芯片以及高通量计算机的若干技术问题,使中国在未来互联网领域的技术竞争中占据主导地位。

3.3 HPC in Cloud

HPC Cloud可能给未来高性能计算市场带来的影响最大。这种基于云计算理念构建的HPC服务,主要面向对计算规模和性能要求较低的中低端HPC用户,在平摊了设备购置和运维成本的同时,向用户屏蔽了复杂的高性能计算机技术细节,降低了高性能计算机的使用门槛。公有云提供商是这一趋势的主要推动者,亚马逊AWS首先推出HPC服务,用户可以创建数千处理器规模的虚拟高性能计算机系统,2011年亚马逊的一台虚拟HPC系统获得了世界Top 500排名的第42位,Linpack效率接近70%。2015年底,阿里云也发布了中国首个云上高性能计算平台。

HPC in Cloud可能成为未来超算中心的重要运营模式之一,若越来越多的HPC用户形成购买HPC服务而非自建系统的习惯,势必形成计算资源的聚

集,未来的中低端 HPC 计算机市场可能被云计算中心所主导。

3.4 新兴使能技术带来新机遇

日益涌现的新兴使能技术,如 3D 堆叠技术、光子学、忆阻器、磁基半导体技术等,对目前仍占主导地位的 CMOS、DRAM 和磁盘等成熟技术提出了挑战。若量子计算、光计算或 DNA 计算等新型计算理论和技术发展成熟,

更将从根基上颠覆整个现代计算机体系。

每一个新兴使能技术孕育着颠覆性创新的机会,这迫切需要体系结构的创新,以挖掘新技术的全部潜力。随着 3D 堆叠等新技术的成熟,很可能会出现面向高性能计算的“大芯片”产品,在单芯片内集成众核 CPU、高速互连以及加速器等,这将带来处理器体系结构上

的重大革新。基于磁基的半导体新工艺速度比传统硅基可提升 1000 倍以上。基于光子学的全光通信技术,可极大降低通信成本。此外,非易失性内存技术(如忆阻器和相变存储器)也驱动着系统设计人员对内存和外部存储系统之间关系的重新思考。因此,新兴使能技术是“弯道超车”的机会,也应该是中国高性能计算领域基础研究的重点。

参考文献 (References)

- [1] Higbie L C. Tutorial: Supercomputer architecture[J]. Computer, 1973, 6(12): 48-58.
- [2] IDC. 2015, IDC HPC Update at ISC'15[EB/OL]. [2016-04-28]. <http://www.slideshare.net/insideHPC/hpc-market-update-from-idc-51061896>.
- [3] Solomon S. Climate change 2007—the physical science basis: Working group I contribution to the fourth assessment report of the IPCC[M]. Cambridge: Cambridge University Press, 2007.
- [4] Kohl S, Leitzl K H, Schmidt M. Transient numerical simulation of CO₂ laser fusion cutting of metal sheets: Simulation model and process dynamics[C]// Proceedings of the 37th MATADOR Conference. Manchester: Springer, 2012: 403.
- [5] Mohamed A, Yu D, Deng L. Investigation of full-sequence training of deep belief networks for speech recognition[C]. INTERSPEECH 2010, 11th Annual Conference of the International Speech Communication Association, Makuhari, Chiba, Japan, September 26-30, 2010.
- [6] Seide F, Li G, Yu D. Conversational speech transcription using context-dependent deep neural networks[C]. INTERSPEECH 2011, 12th Annual Conference of the International Speech Communication Association, Florence, Italy, August 27-31, 2011.
- [7] The New York Times. How many computers to identify a cat? 16,000[J/OL]. Communications of the Acm, 2012-06-26[2016-04-28], <http://www.cs.cornell.edu/courses/CS6700/2013sp/readings/04-a-Deep-Learning-NYT.pdf>.
- [8] Chi K R. Neural modelling: Abstractions of the mind[J]. Nature, 2016, 531(7592): S16-S17.
- [9] Moore G. Moore's law[J]. Electronics Magazine, 1965, 38(8): 114.
- [10] The Top500 List 2015. [EB/OL]. [2016-04-20]. Available: <http://www.Top500.org>.
- [11] Bryant R E, Polk W T. The National Strategic Computing Initiative[C]//Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis: ACM, Austin 2015.
- [12] Murphy R C, Wheeler K B, Barrett B W, et al. Introducing the graph 500[J]. Cray User's Group (CUG), 2010, 14(3): 15555-15558.
- [13] The Green500 List. [EB/OL]. [2016-04-22]. Available: <http://www.Green500.org>.

The development of high-performance computing

ZANG Dawei, CAO Zheng, SUN Ninghui

Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China

Abstract Following theoretical science and experimental science, high-performance computing has become the third largest paradigm of human exploration. As an important means of technology innovation, high-performance computing is widely used in many areas, such as nuclear explosion simulation, climate simulation, and engineering calculations. Moreover, it is a strategic commanding point of modern sci-tech competition and embodies the comprehensive strength of a nation. In this paper, we describe the history and current situation of high-performance computing, then analyzes the problems and challenges. Also, we discuss the future direction of high-performance computing.

Keywords high-performance computing; high-performance computer; high-performance computing market

(编辑 傅雪)