

戈登奖^{*}

——分析与思考

Review of the Gordon Bell Prize

张理论¹, 邓小刚^{1,2}

ZHANG Li-lun¹, DENG Xiao-gang^{1,2}

(1. 国防科学技术大学计算机学院, 湖南 长沙 410073;

2. 中国空气动力研究与发展中心空气动力学国家重点实验室, 四川 绵阳 621000)

(1. School of Computer Science, National University of Defense Technology, Changsha 410073;

2. State Key Laboratory of Aerodynamics,

China Aerodynamics Research and Development Center, Mianyang 621000, China)

摘 要:戈登奖(Gordon Bell Prize)是高性能计算应用领域的最高学术奖项。与 TOP500 重视衡量高性能计算机系统性能相比,该奖项更关注用于解决重要科学问题的高性能计算技术创新,是国际上公认的高性能计算应用技术发展水平的重要标杆。本文综合分析近年来戈登奖的获奖研究成果,尤其是最高性能奖和特别奖的研究特点及科学意义。在此基础上,总结规律,并就如何推进高性能计算应用研究,给出了一些思考,以期为我国从事超级计算应用研究的同仁提供参考。

Abstract: The ACM Gordon Bell Prize is the top academic award in the field of HPC application. This prize is awarded to recognize outstanding achievement, with particular emphasis on innovation in applying HPC to applications in science. Now it is deemed to be the measure of the progress over the time of HPC application. In this paper, the results of the award-winning research are summarized and analyzed over the last thirteen years, especially for the peak performance and the special awards. Some views on how to promote the research of HPC application are presented for the reference of demestic colleagues.

关键词:高性能计算;科学应用;戈登奖;最高性能奖;特别奖

Key words: HPC; scientific application; Gordon Bell Prize; peak performance award; special award

doi:10.3969/j.issn.1007-130X.2012.08.006

中图分类号: TP338.6

文献标识码: A

1 引言

戈登奖(Gordon Bell Prize)是由 VAX 系列机的先驱者一贝尔·戈登于 1987 年倡议发起的,从每年超级计算大会的专题论文中筛选、评审产生,并由 ACM 颁发,是超级计算应用领域的国际最高学术奖。戈登奖的产生有一个前奏。上个世纪 80

年代,囿于 Amdahl 定律关于并行程序中串行部分的比例限制加速比的认识,学术界对并行计算能否获得高加速比感到困惑。1986 年,Alan Karp 博士提议:谁先在通用计算机上针对实际应用获得超过 200 倍的加速比,他将支付 100 美元作为奖励。1987 年,第一届戈登奖获得者以在 1 024 个处理器上的 400 倍加速比同时也获得了 Karp 奖。获奖者中有后来在高性能计算界广为人知的 John L

^{*} 收稿日期:2012-04-28;修订日期:2012-06-11
通讯地址:410073 湖南省长沙市国防科学技术大学计算机学院软件所
Address: Institute of Software, School of Computer Science, National University of Defense Technology, Changsha, Hunan 410073, P. R. China

Gustafon,他提出的 Gustafson 定律^[1]为并行算法研究和应用注入了新的活力。

戈登奖重视超级计算的实际应用。据 2012 年超级计算会议主页,“颁发戈登奖是为了肯定高性能计算的杰出成就、追踪并行计算进展,重点关注将高性能计算用于科学问题的有益创新”。奖项主要分为三类:最高性能奖、特别奖和最佳性价比奖。最高性能奖授予达到最高持续浮点性能的实际并行应用。最佳性价比奖授予体现最佳性价比的并行应用及其平台。特别奖的获奖应用虽未达到最高性能,但通过技术创新达到了新的性能水平,这些创新涉及问题建模、算法、数据结构、并行实现和优化等能促进超级计算应用进步的各个环节。有的特别奖在名称上添加个性色彩以反映其特点,如 2002 年的“编程语言”(Language)、2003 年的“系统生命周期”(Lifetime)和 2008 年的“算法创新”(Algorithmic Innovation)等。

2 戈登奖研究内容概要与综合分析

2.1 戈登奖研究内容概要

2.1.1 最高性能奖

1996 年以前的获奖研究多采用早期的共享存储向量计算机。1996 年以来,1996 年、1999 年、2002 年、2003 年、2004 年的获奖研究均基于偏微分方程(Partial Differential Equation,简称 PDE)模型,以计算流体力学(Computational Fluid Dynamics,简称 CFD)应用为主。1996 年获奖的研究“数值风洞(Numerical Wind Tunnel,简称 NWT)计算机上的流体动力学模拟”^[2]是由日本国家航空宇航实验室完成的。数值风洞计算机是当时日本主要用于 CFD 的通用计算机,共有 166 个向量处理器,理论峰值性能为 0.282 万亿次/秒。该应用模拟空气以一定角度进入飞机引擎时的气流状态,此时引擎转子叶片受力不均匀,极端情况下会导致转子停转,飞机失速。对此设计了一种静态和动态网格功能相结合的混合网格系统,在 160 个处理器上获得 0.111 万亿次/秒的持续性能。1999 年获奖研究为“在 IBM-SP 系统上的高分辨率可压湍流模拟”^[3],该研究由美国能源部劳伦斯利弗莫尔实验室、明尼苏达大学和 IBM 公司合作完成。研究采用高精度分段抛物方法(Piecewise Parabolic Method,简称 PPM),模拟了激波穿越两种流体交界面的情形,这是化学反应流和燃烧中的典型现象。在 IBM RS/6000 计算平台上,利用 240 亿单元的数据规模进行了原理

验证,在 5 832 个 CPU 上达到 1.18 万亿次/秒的持续性能;采用 80 亿单元的数据规模模拟了激波管实验,在 3 840 个 CPU 上运行墙钟时间超过 1 周,持续速度达到 0.6 万亿次/秒(32 位精度)。对此,Bill Dannevik 给出的评价为“湍流模拟达到了前所未有的高分辨率;结合实验室辅助或流场实验,有潜力获得系列科学突破”;Ron Cohen 的评价为“如此高的分辨率,使研究运动尺度差异显著的流体之间的相互作用及其反应机制成为可能”。2000 年、2001 年的获奖研究均为基于专用计算机 GRAPE-6 系统的天体物理数值模拟^[4,5]。2002 年、2003 年、2004 年的获奖研究平台为地球模拟器(Earth Simulator),研究领域分别为气候、地震、地磁。2002 年获奖研究为“地球模拟器上基于谱变换方法的全球大气环流 26.58 万亿次数值模拟”^[6],其主要贡献为:采用水平分辨率 10 公里地球模拟器全球大气环流模式(AGCM for Earth Simulator,简称 AFES),在全系统 5 120 个向量处理器上达到 26.58 万亿次/秒的持续性能,计算效率为 64.9%。该应用展示了地球模拟器解决复杂实际应用问题的能力。2003 年获奖研究为“地球模拟器上 5 万亿次持续速度、2.5TB 存储和 146 亿自由度的地震数值模拟”^[7]。基于震波传播模拟软件 SPEC-FEM3D,考虑了真实地震模拟所需的三维波传播、结构密度、地球曲率、地形拓扑和海洋地形等因素,利用 1 944 个 CPU(全系统的 38%),计算效率为 30%。2004 年获奖研究为“地球模拟器上 15.2 万亿次持续速度的地磁电流数值模拟”^[8]。一般认为,地磁是由地球内部高温高压的金属流体因地球自转产生的自激电流所导致,该研究利用 4 096 个向量处理器,获得每秒 15.2 万亿次/秒的持续速度,计算效率为 46%。

2005 年获奖研究为“BlueGene/L 系统上的超过百万亿次持续性能的(金属)凝固数值模拟”^[9]。采用研制的分子动力学代码 ddcMD,首次从原子层次研究了金属钽和铪在超高温高压下的凝固过程。在 IBM BlueGene/L 全系统的 131 072 个 CPU 上,达到 101.7 万亿次/秒的持续速度。据称该研究对评估核武器储备的可靠性具有重要意义。2006 年获奖研究为“BlueGene/L 系统上的大规模重金属结构计算”^[10]。采用第一原理分子动力学(First Principles Molecular Dynamics,简称 FPMD)代码 Qbox 计算金属钼结构,在 IBM BlueGene/L 全系统上达到 207.3 万亿次/秒的持续速度,计算效率为 56.5%。据称该数值模拟对评估高温高压下重金属的结构和电子状态具有重要意义。另外,2006 年还有一项提名奖^[10],采用专用计

算机 MD-GRAPE3。2007 年获奖研究为“增强计算稳定性,降伏‘CPU 千年虫’—开尔文·赫姆霍茨(K-H)不稳定性的微米尺度原子级模拟”^[12]。该应用是第一个基于分子动力学的微米尺度原子级 K-H 不稳定性数值模拟,采用 ddcMD 代码在 IBM BlueGene/L 全系统上研究了二元合金熔化时在界面 K-H 不稳定性的发生发展过程,获得了 54.3 万亿次/秒的持续性能。该研究揭示了 K-H 不稳定性从原子尺度向宏观连续动力学尺度的演变过程。2008 年获奖研究为“采用新算法使高温超导失序效应计算获得超过 400 TFlop/s 持续性能”^[13],分析不同物质的库珀电子对在不同温度下的形成发展过程。数值模拟软件 DCA++ 经橡树岭国家实验室高性能计算组在访存和混合精度计算两方面的优化,性能提高了约 10 倍。利用 Cray XT4 系统的 4.9 万个 CPU,达到 409 万亿次/秒(混合精度)的持续速度。2009 年获奖研究为“纳米系统自由能的可扩展计算方法”^[14]。采用软件 WL-LSMS,在美洲虎(Cray XT5)全系统上(223 000 CPU 核)获得 1 840 万亿次/秒的持续速度。该研究准确计算了磁体失去磁性的温度条件,为研制稳定的磁性材料提供了科学依据,对设计新型存储介质具有重要参考价值。

2010 年获奖研究为“在 20 万 CPU 核和异构体系结构上的千万亿次持续性能血流模拟”^[15]。血液流动是一个多物理、多尺度问题,其数值模拟主要涉及血浆流体流动计算和(变形)红细胞粒子模拟,该研究采用面向血液微循环直接数值模拟(Direct Numerical Simulation,简称 DNS)的软件框架 MOBO,同时利用 PETSc 作为性能优化工具,基于 DENDRO 软件构造树结构,进行红细胞数据重剖分。在三个平台上(分别是戴尔机群、美洲虎和一个 NVIDIA Fermi GPU 结点)进行测试。其中,在美洲虎的 19.6 万个 CPU 核上达到 780 万亿次/秒的持续性能。该研究的分辨率比以往提高了上万倍,使科学家能更加真实地观察血液流动状态,更好地理解血栓形成过程,从而改进微射流控制设备的设计,对血栓的早期病理学诊断及抗血栓药物的研制具有潜在价值。

2011 年获奖研究为“在 K 计算机上 10 万原子规模、基于第一原理的硅纳米丝电子结构数值模拟”^[16],13 位作者来自于日本理化研究所(RIKEN)、富士通公司等。硅纳米丝作为下一代场效应晶体管的导体材料,由 1 万~10 万个硅原子组成,直径 10~20 纳米,长度约 10 纳米。该研究首次针对

包含 107 292 个硅原子、长度为 6 纳米、直径为 20 纳米的硅纳米丝,基于量子理论计算了其电子状态。利用 K 系统在理化研究所的后期调试阶段开展研究,采用的软件名为 RSDFT。研究人员针对 K 系统的专用互连结构,优化了通信库,将并行任务对网络拓扑做了近似映射。采用 442 368 个 CPU 核,达到了 3 080 万亿次/秒的持续性能,计算效率为 43.6%。数值模拟显示,硅纳米丝形状对电子输运特性存在很大影响。

2.1.2 特别奖

1999 年获奖研究为“非结构网格 CFD 应用获得高持续性能”^[17],5 位来自 NASA 兰利研究中心、阿肯实验室和弗杰尼亚大学的作者,将兰利研究中心的非结构网格有限体积 CFD 软件 FUN3D(该软件可用于飞机、汽车和潜艇的外形设计)移植到 PETSc 并行计算框架上,并基于 PETSc 的 Newton Krylov 非线性解算器和 Schwarz 预条件,开展了数据交错、结构块和顶点重排序等一系列以提高访存性能为核心的优化技术。利用 IBM ASCI-Red 的 2 048 个双 CPU 结点,对数据规模为 280 万单元的定常欧拉流进行数值模拟,达到 0.156 万亿次/秒的持续速度。此前,1988 年、1989 年、1990 年与 1996 年的最高性能奖均基于结构网格。该研究展示了非结构网格 CFD 隐式求解的高并行可扩展性。2000 年获奖研究为“在数千个 CPU 上基于自适应网格细分的高性能反应流数值模拟”^[18],13 位作者分别来自芝加哥大学、劳伦斯利弗摩尔实验室、Intel 公司等 5 个单位。该研究采用了天体物理领域的可压反应流软件 FLASH 和块结构自适应网格细分软件 PARAMESH,对 PARAMESH 进行了 MPI 通信优化,对 FLASH 进行了计算方法优化。在 IBM ASCI-Red 计算机上,利用 6 420 个 CPU 获得了 0.238 万亿次/秒的持续性能,实现了“最大规模和最高空间分辨率的超新星爆炸模拟”。

2001 年特别奖为“基于 Cactus 并行框架和 Globus 网格工具集的高效异构分布并行环境”^[19],基于 Cactus 和 Globus 两个软件,构造了一个网格计算支撑环境。其主要贡献为:采用 Globus 的 MPI 实现—MPICH-G2,使 Cactus 并行框架具备网格计算支持能力;利用来自 4 台不同计算机的共计 1 500 个处理器实现了爱因斯坦方程求解的异构并行。2002 年特别奖的具体名称是“编程语言奖”(Award for Language),获奖论文是“地球模拟器上基于 HPF 语言的 14.9 万亿次核聚

变三维流体数值模拟”^[20]。IMPACT-3D 是一个面向三维可压、无粘欧拉流的等离子体数值模拟软件,主要用于核装置内爆分析。该研究基于 HPF 编程语言实现了 IMPACT-3D 程序的向量并行化。利用地球模拟器的 512 个节点(4 096 个向量处理器),通过手工调节通信优化,对 $2048 \times 2048 \times 4096$ 离散规模问题,获得了 14.9 万亿次/秒的持续速度,计算效率为 45%。

2002 年还有其他三个特别奖。(1)“地球模拟器上基于谱变换方法的、达到 16.4 万亿次持续性能的湍流数值模拟”^[21]。采用新方法实现三维快速傅立叶变换,克服了湍流数值模拟的主要计算瓶颈。(2)“Salinas: 一个可扩展结构力学数值模拟软件”^[22]。Salinas 是一个面向复杂结构的有限元结构力学分析软件,兼具静态和动态有限元分析功能。利用 IBM ASCI White 计算机的 3 375 个 CPU,问题自由度超过 1 亿,获得 1.16 万亿次/秒的持续性能。(3)“NAMD: 数千个处理器上的分子生物学数值模拟”^[23]。NAMD 是一个全功能、产品级的分子生物学软件,基于面向对象语言 Charm++, 利用 Elan 并行通信库替代了 MPI,该应用在 Alpha 机群 Lemieux 上,扩展至 2 250 个 CPU 规模。

2003 年有两项特别奖。(1)“万亿次系统上的高分辨率正向和逆向地震模型”^[24]。对复杂三维地质构造,如何确定地震震源、变形和衰减参数等是具有挑战性的反问题。对此,该研究基于六面体网格发展了可扩展至 10 亿网格的震波传播求解器。对 1 亿网格规模问题,利用 3 000 个 Alpha Server 处理器获得了 1.21 万亿次/秒的持续性能,并行效率为 80%。(2)“面向 40 万亿次实际性能的 GRAPE-6 系统性能评估与调优”^[25]。GRAPE-6 系统含 2 048 个定制流水处理器,理论峰值为 63.4 万亿次/秒,在 Benchmark 测试中,最高持续性能达到 35.3 万亿次。奖项鼓励 GRAPE-6 系统较长的使用生命周期。

2004 年特别奖为“5 亿以上自由度的超可扩展有限元固体力学分析”^[26],该研究基于 IBM ACSI White 计算机,开展人体脊柱的有限元弹性结构力学分析;利用并行有限元软件 Athena、非结构网格剖分软件 ParMetis 和 PETSc 并行框架,在 4 088 个处理器上进行 5.37 亿自由度的数值模拟。其中,椭圆问题求解采用基于代数多重网格的非精确牛顿迭代方法,据称是当时并行规模最大的复杂几何变形、非结构有限元弹性力学数值模拟。2006 年特别奖为“BlueGene/L 系统与量子色动力学”^[27],采用 32

000 个 CPU,持续性能为 12.2 万亿次/秒,计算效率为 20%。2005 年、2007 年特别奖空缺。2008 年奖项称为“算法创新奖”(Prize for Algorithmic Innovation),获奖论文为“面向数千个原子纳米尺度结构的线性可扩展、分而治之的电子结构计算”^[28]。电子结构计算研究往往涉及 1 000~100 000 个原子,已有的 $O(N)$ 复杂度快速方法存在严重的并行通信瓶颈,难以扩展至 10 000 个 CPU 核,原子规模不超过 2 000 个。对此,该研究提出了一种新的线性可扩展三维片段方法(Linearly Scaling Three-Dimensional Fragment,简称 LS3DF),在保持 $O(N)$ 复杂度和计算精度的同时实现了极小通信。在 BlueGene/P 全系统的 131 072 个 CPU 核上,对 13 842 个原子规模问题,获得 107.5 万亿次/秒的持续速度,计算效率为 24.2%。该研究对设计新型太阳能电池材料具有重要价值。

2009 年有两项特别奖。(1)“呼之欲出的猫:包含 10^9 规模神经元、 10^{13} 规模突触的大脑皮质模拟”^[29]。该认知计算研究受到美国国防部先进研究计划署的支持,在 IBM BlueGene/P 上,利用 147 456 个 CPU 核及 144TB 内存开展了大规模大脑皮质功能模拟,其神经元和突触规模与猫大脑相当。(2)“Anton 计算机上的毫秒级分子动力学数值模拟”^[30]。一般来说,生物分子动力学系统数值模拟的典型步长为 10^{-15} 秒,而能刻画生物分子变化过程的门槛时间为毫秒到微秒级。采用专用计算机 Anton 首次开展为期 1 毫秒的生物分子动力学数值模拟,是以往模拟时间长度的 100 倍以上。2010 年特别奖为荣誉提名,具体名称为“可扩展性特殊成就-荣誉提名”(Honorable Mention for Special Achievements in Scalability),获奖论文为“纳米材料科学中激发态和强相关的第一原理电子结构数值模拟”^[31]。

2011 年特别奖为“TSUBAME 2.0 系统上树状凝固的千万亿次相场模拟”^[32]。奖项具体名称为“可扩展性和求解加速的特殊成就(Special Achievement in Scalability and Time-to-Solution)”。相场方法是研究二元合金微尺度凝固过程的重要手段,由于涉及到细网格上复杂非线性项计算,传统三维凝固过程模拟仅限于简单形状。研究平台 TSUBAME 2.0 系统共有 1 408 个 HP Proliant SL390s 结点,单结点含两个 Intel CPU 和三个 Tesla M2050 GPU。该研究工作采用了 4 000 个 GPU 和 16 000 个 CPU 核,持续性能为 1 017 万亿次/秒(单精度)。据称这是目前唯一的持续速度超过千万亿次的基于网格离散的数值模拟问题。

2011 年还有三项荣誉提名奖。(1)“持续速度达到 1 440 万亿次的原子级纳米电子设备工程学”^[33]。研究了不同几何形状和材质条件下的纳米级晶体管性能,所采用的软件 OMEN 是下一代纳米电子设备的 CAD 工具;利用美洲虎计算机的 221 400 个 CPU 核,获得 1 440 万亿次/秒的持续速度,计算效率为 55%。据称该研究开启了纳米晶体管设计的新篇章。(2)“200 万核上的千万亿次生物流体模拟”^[34]。采用多物理多尺度代码 MUPHY,该代码基于面向流体的 Lattice Boltzmann 方法和针对悬浮体的分子动力学方法及其耦合;主要工作体现在复杂几何区域的数据剖分、流体动力系统与粒子动力系统间的通信、流体和粒子系统间的相互作用模型、CPU 在 GPU 间发挥高效通信协处理器作用等;采用 TSUBAME 2.0 系统的 4 000 个 GPU 获得每秒 600 万亿次/秒的持续性能。据称这是第一个基于 X 射线血管造影数据的冠状动脉心血管血流大规模多尺度数值模拟。(3)“多尺度数值模拟的一个新范例:大脑血流应用”^[35]。在宏观尺度上采用非定常高阶谱有限元软件 NekTar 求解 NS 方程;在微观尺度采用分子动力学软件 DPD-LAMMPS。其主要贡献包括数值模拟重叠区的界面模型、互连拓扑感知的多层通信、多尺度可视化以及 SIMD 优化为核心的单机优化等;分别采用 IBM BlueGene/P 和 Cray XT5 进行并行计算;最大并行规模达到 190 740 个处理器核(Cray XT5)。据称这是第一个针对实际生物系统的大规模多尺度数值模拟及可视化实现。

2.1.3 最佳性价比奖

最佳性价比奖的衡量标准为:在单位时间每获得 1M 浮点性能需要投入的资金(单位: Dollar per MFlops)或者每投入 1 美元能够获得的浮点性能。这里统一称为性价比。该奖项的意义在于,针对特定的实际应用,寻求软硬件资源的最优化配置。表 1 给出了 1999 年以来天体物理 N 体问题应用的最佳性价比奖简况^[36~39]。

表 1 最佳性价比奖情况

年份	平台	dollars per MFlops
1999	GRAPE5	7.0
2006	CRAY	0.185
2009	GPU 集群	0.008
2010	GPU 集群	0.004

2.2 综合分析

表 2 给出了 1988 年以来的最高性能奖应用的综合情况。自 1988 年以来,除 1990 年、1991 年外,

最高性能奖每年常设。在 2004 年及以前的共计 15 次最高性能奖获奖应用中,8 次是以偏微分方程为模型基础,有 4 次(1996 年、1999 年、2002 年、2004 年)属于 CFD 研究范畴。2010 年的血液流动是包含 CFD 与分子动力学的多物理问题。2005 年及以后的应用主要以分子动力学、量子力学、蒙特卡罗及其结合为模型基础,涉及核材料、纳米材料和生物信息学等,且有向多尺度、多物理方向发展的特征。鉴于 PDE 求解在现代科学研究和工程实际中的地位至关重要,加上分子动力学、蒙特卡罗等模型通常比较适合大规模并行,使人怀疑戈登奖是否有了某种程度的“唯持续性能”倾向。无论如何,戈登奖的获奖研究的确代表了现代科学前沿问题数值模拟的最高水准,集中体现了“现代物质科学向微观深入、向宏观拓展、向极端条件发展”^[40]的大趋势。

最高性能奖代表了实际应用问题利用同期高性能计算机的最大能力。图 1 显示应用问题的最高持续性能总体上呈上升趋势,但与 TOP500 的 Linpack 测试最高性能^[41]之间往往存在一定差距。但是,值得注意的是,戈登奖最高持续性能增长也呈“每十年大约增长一千倍”的速度,这与高性能计算机的性能增长速度类似。一方面,应用对系统的这种“追赶式”发展体现了高性能计算系统技术进步为开展大规模应用研究提供了平台和可能。另一方面,在多数情况下,应用的最高持续性能可以达到或超越 2~3 年前的 Linpack 性能,这要归因于应用领域数学物理模型、方法和软件技术的不断进步。当然,这也从应用的角度肯定了 Linpack 性能测试的合理性和重要参考价值。由于不同应用的数学模型、离散方法及实现手段等对并行可扩展性的限制不一,不同领域应用在同一历史时期所能达到的最高持续性能存在显著差别,因而历年来最高性能奖的持续性能并非严格递增,在一些年份明显下降,如 2003 年、2007 年和 2010 年。高性能计算应用在领域间的这种不均衡发展格局,也凸显了开展建模和计算方法研究的重要性。

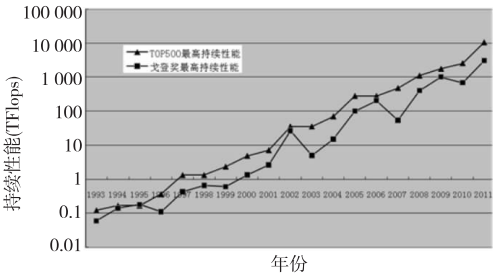


图 1 历年最高性能奖的持续性能和 TOP500 最高持续性能

表 2 最高性能奖情况综合(持续性能单位:TFlops)

年份	模型	问题类型或者领域	CPU(核)数	持续性能	计算平台
1988	偏微分方程	静态有限元	8	0.001	Cray Y-MP
1989	偏微分方程	震波传播		0.006	CM-2
1992	N 体问题	万有引力、天体物理		0.005	Intel Touchstone Delta
1993	偏微分方程	CFD、激波计算	1 024	0.06	CM-5
1994	偏微分方程	边界元、结构力学	1 904	0.14	Intel Paragon
1995	量子色动力学	夸克、基本粒子研究	128	0.179	Numerical Wind Tunnel
1996	偏微分方程	CFD、气动计算	166	0.111	Numerical Wind Tunnel
1997	N 体问题	万有引力、天体物理	4 096	0.43	ASCI Red
1998	量子分子动力学	原子排列、金属磁性	1 024	0.657	Cray T3E
1999	偏微分方程	CFD、湍流	3 840	0.6	Blue Pacific
2000	N 体问题	万有引力、天体物理	96	1.349	GRAPE-6
2001	N 体问题	万有引力、天体物理	1 024	11.55	GRAPE-6
2002	偏微分方程	CFD、大气环流	5 120	26.58	Earth Simulator
2003	偏微分方程	震波传播	1 944	5	Earth Simulator
2004	偏微分方程	CFD、磁流体	4 096	15.2	Earth Simulator
2005	分子动力学	凝固、核材料	131 072	101.7	BlueGene/L
2006	量子分子动力学	电子结构、核材料	131 072	207	BlueGene/L
2007	分子动力学	K-H 不稳定性、核材料	131 072	54.4	BlueGene/L
2008	量子蒙特卡罗	电子结构、超导材料	49 288	409	Cray XT4
2009	量子蒙特卡罗	纳米热动力学、磁性材料	147 464	1 030	Cray XT5
2010	多物理	CFD+分子动力学、生物医学	196 608	700	Cray XT5
2011	量子分子动力学	电子结构、半导体材料	442 368	3 080	K 系统

图 2 反映了最高性能奖往往是跨学科、跨单位合作研究的成果(特别奖类似,下略)。在 13 年中,有 12 年是两个及以上单位合作,7 年为三个及以上单位合作,4 年为四个以上单位合作。这些合作单位通常包括国际知名的实验室、研究中心、大学或大公司,可以分为应用问题提出单位(数据与验证)、数学物理模型和计算方法研究单位、高性能计算机研制单位三类(有时一个单位同时具备两种以上特征)。

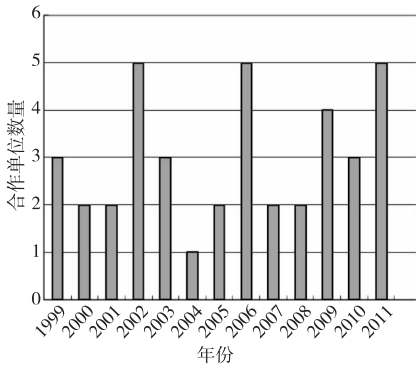


图 2 最高性能奖的合作研究情况

图 3 显示,特别奖涉及的领域与最高性能奖相比,大体上是类似的。特别奖的技术创新多种多样,侧重点覆盖高性能计算应用的问题模型、算法、编程语言、并行框架、性能优化等多个层面,其中强调应用并行可扩展性、应用模型和算法突破的居一

半以上。

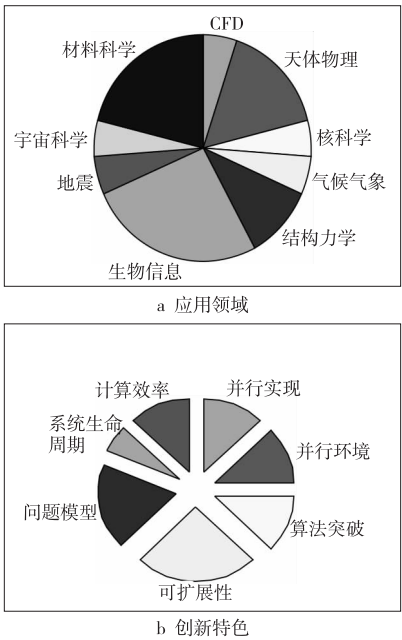


图 3 历年特别奖的应用领域与研究特色

最佳性价比奖是两个部分的优化组合,即高性能计算系统价格与应用在系统上的持续性能。图 4 显示,对天体物理数值模拟而言,获得 1MFlops 性能的价格由 1999 年的 7 美元降至 2010 年的 0.004 美元,平均速率超过摩尔定律。这是由于最佳性价比奖包含了系统和应用两个层面的技术进

步及其结合,而摩尔定律仅是对系统器件发展的概括。

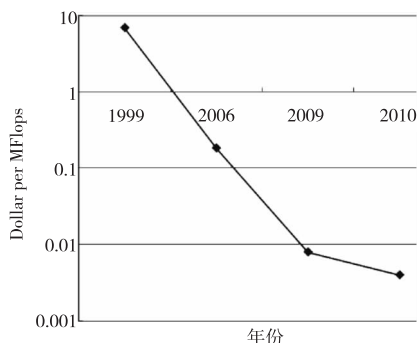


图4 天体物理应用获最佳性价比奖的指标变化

最高性能奖和特别奖的获奖平台几乎都是同期 TOP500 排名前三的系统,这反映了高端计算机系统是催生重大科学发现的温床。另外,2009 年最高性能奖涉及磁性材料研究,2011 年最高性能奖和特别奖涉及新型纳米半导体材料研究,分别与新型存储介质和纳米器件等计算机基础使能技术息息相关,这也说明大规模并行应用研究,反过来会对计算机系统技术自身的发展产生重要的影响。

3 几点思考

戈登奖从科学研究的角度凸显了高性能计算应用的发展方向,即更加贴近物理真实,这就提出了数学物理模型更合理、高时空分辨率、计算方法高精度和高并行可扩展性等一系列要求。结合全文,就如何推进高性能计算应用研究,尝试给出以下几点思考:

(1)从应用领域自身看来,领域科学研究的积累和较高的水准,是利用高性能计算机实现研究水平再上台阶的前提。应用问题的科学与工程意义是开展大规模并行数值模拟的基本出发点,这要求数学物理模型、方法及软件达到相当的成熟度,且具备利用超级计算机开展数值模拟所需的真实数据和验证条件。

(2)从应用领域与计算机、数学等学科之间的关系看,需要实现多应用领域、计算机科学、数学等多学科交叉协同,模型、算法、软件实现和性能优化等全方位结合。绝大多数的获奖研究团队由两个及以上单位组成,跨单位、跨学科专业的特征显著,这种多学科团队的组成是高性能计算应用研究的内在要求,多尺度、多物理真实科学问题的极大规模高效并行计算是问题模型、算法、并行实现乃至

鲁棒性等多因素综合的结果。

(3)从研究平台的角度看,需要重视高端计算平台和新型体系结构系统的应用。迄今为止,戈登奖一直被美、日两国包揽,其研究平台通常就是所在国的高端计算机。近几年来,出现了数个基于 GPU/CPU 异构平台的获奖应用,这些研究工作也推动了异构混合并行算法和软件的发展。

(4)从建模和计算方法对性能限制的角度看,要注重模型和数值方法的创新研究。从模型和算法层面创新是高性能计算应用获得重大进展的基础性工作。目前,尚无真正达到千万亿次持续性能(64 位)的 PDE 求解类实际应用。在极大规模并行时代,传统的 PDE 求解面临并行计算瓶颈,这主要归因于离散格式和算法所导致的高通信计算比和高访存计算比。发展极小通信和低访存计算比的模型和数值算法,或将成为需求旺盛的研究方向。

(5)从应用软件角度看,要注重继承和发展已有的数值并行计算软件成果。大多数戈登奖获奖应用的核心软件是“in house”代码,是对多年研究成果的继承,甚至还出现了 2005 年和 2007 年的最高性能奖基于同一个代码—ddcMD 的情况。除这些“in house”代码外,一些应用部分采用了开源软件作为性能支撑工具,例如 1999 年特别奖、2004 年特别奖、2010 年最高性能奖采用了 PETSc 数值并行软件框架,2001 年特别奖采用了 Cactus 并行软件框架。

(6)从数值模拟的输入输出数据角度看,需要重视数据生成、存储和可视化,将其与模型算法研究、软件实现、性能优化等加以统一考虑。在超高的时空分辨率下,数值模拟所需要的初始数据生成和结果数据存储将会成为瓶颈。传统上采用串行方法生成网格数据的做法,在离散规模达到数百亿、千亿甚至更多时需要考虑并行。另外,海量结果数据的生成、存储和并行可视化同样需要加以重视。

4 结束语

我国的“天河一号”系统首次将 CPU+GPU 异构融合体系结构成功用于高端超级计算机,获得 2010 年的 TOP500 第 1 名,引领了国际高性能计算机系统技术的发展,为超级计算发展史增添了中国的浓墨重彩。反观超级计算应用,相对国际水平,我国在深度和广度方面还存在较大的差距。高

性能计算应用研究是一项集多领域、学科大成的科学研究活动,是问题模型、离散方法、验证与确认、快速并行算法、并行软件设计、性能优化和可视化等涉及领域科学、计算机科学和数学等多位一体的系统性工作。通过物理、数学和计算机多学科交叉,开展协同创新和融合式研究,从领域科学问题的实际需求出发,从建模、算法、软件、前后处理及与计算机系统的有效结合等多个层面开展针对性强的工作,才能确保极大规模数值模拟的重要科学意义,解决好负载平衡、可扩展性、鲁棒性以及大数据前后处理等一系列瓶颈问题。

参考文献:

- [1] Gustafson J L, Montry G R, Benner R E. Development of Parallel Methods for a 1024-Processor Hypercube [J]. SI-AM Journal on Scientific and Statistical Computing, 1988, 9 (4): 609-638.
- [2] Karp A H, Geist A, Bailey D. 1996 Gordon Bell Prize Winners[J]. IEEE Computer, 1997, 30(1): 80-85.
- [3] Mirin A A, Cohen R H, Curtis B C, et al. Very High Resolution Simulation of Compressible Turbulence on the IBM-SP System [C] // Proc of the 1999 ACM/IEEE Conference on Supercomputing, 1999: 70.
- [4] Makino J, Fukushima T, Koga M. A 1.349 TFlops Simulation of Black Holes in a Galactic Center on GRAPE-6 [C] // Proc of the ACM/IEEE SC2000, 2000: 43.
- [5] Fukushima T, Makino J. 11.5 TFlops Simulation of Black Holes in a Galactic Center on GRAPE-6 [C] // Proc of the ACM/IEEE SC'01, 2001.
- [6] Shingu S, Takahara H, Fuchigami H, et al. A 26.58 TFlops Global Atmospheric Simulation with the Spectral Transform Method on the Earth Simulator [C] // Proc of the ACM/IEEE SC'02, 2002.
- [7] Komatitsch D, Tsuboi S, Ji C, et al. A 14.6 Billion Degrees of Freedom, 5 Teraflops, 2.5 Terabyte Earthquake Simulation on the Earth Simulator [C] // Proc of the ACM/IEEE SC'03, 2003.
- [8] Kageyama A, Kameyama M, Fujihara S, et al. A 15.2 TFlops Simulation of Geodynamo on the Earth Simulator [C] // Proc of the ACM/IEEE SC'04, 2004: 35.
- [9] Streitz F H, Glosli J N, Patel M V, et al. 100+ TFlop Solidification Simulations on BlueGene/L [C] // Proc of the ACM/IEEE SC'05, 2005.
- [10] Gygi F, Draeger E W, Schulz M, et al. Large-scale Electronic Structure Calculations of High-z Metals on the BlueGene/L Platform [C] // Proc of the ACM/IEEE SC'06, 2006: 45.
- [11] Narumi T, Ohno Y, Okimoto N, et al. A 185 TFlop/s Simulation of Amyloid-forming Peptides from Yeast Prion Sup35 with the Special-Purpose Computer System MD-GRAPE3 [C] // Proc of the ACM/IEEE SC'06, 2006.
- [12] Glosli J N, Caspersen K J, Gunnels J A, et al. Extending Stability Beyond CPU Millennium: A Micron Scale Atomistic Simulation of Kelvin-Helmholtz Instability [C] // Proc of the ACM/IEEE SC'07, 2007: 58.
- [13] Alvarez G, Summers M S, Maxwell D E, et al. New Algorithm to Enable 400+ TFlop/s Sustained Performance in Simulations of Disorder Effects in High- T_c Superconductors [C] // Proc of the ACM/IEEE SC'08, 2008: 61.
- [14] Eisenbach M, Zhou C G, Nicholson D M, et al. A Scalable Method for Ab Initio Computation of Free Energies in Nanoscale Systems [C] // Proc of the ACM/IEEE SC'09, 2009: 64.
- [15] Rahimian A, Lashuk I, Veerapaneni S, et al. Petascale Direct Numerical Simulation of Blood Flow on 200K Cores and Heterogeneous Architectures [C] // Proc of the ACM/IEEE SC'10, 2010.
- [16] Hasegawa Y, Iwata J I, Tsuji M, et al. First-principles Calculations of Electron States of a Silicon Nanowire with 100,000 Atoms on the K Computer [C] // Proc of SC'11, 2011.
- [17] Anderson W K, Gropp W D, Kaushik D. K, et al. Achieving High Sustained Performance in an Unstructured Mesh CFD Application [C] // Proc of the ACM/IEEE SC'99, 1999: 69.
- [18] Calder A C, Curtis B C, Henry G, et al. High-Performance Reactive Fluid Flow Simulations Using Adaptive Mesh Refinement on Thousands of Processors [C] // Proc of the ACM/IEEE SC2000, 2000: 56.
- [19] Allen G, Dramlitsch T, Foster I, et al. Supporting Efficient Execution in the Heterogeneous Distributed Computing Environments with Cactus and Globus [C] // Proc of the ACM/IEEE SC'01, 2001: 52.
- [20] Sakagami H, Murai H, Seo Y, et al. 14.9 TFLOPS Three-dimensional Fluid Simulation for Fusion Science with HPF on the Earth Simulator [C] // Proc of the ACM/IEEE SC'02, 2002: 51.
- [21] Yokokawa M, Itakura K, Uno A, et al. 16.4 Tflops Direct Numerical Simulation of Turbulence by Fourier Spectral Method on the Earth Simulator [C] // Proc of the ACM/IEEE SC'02, 2002: 50.
- [22] Bhardwaj M M, Pierson K, Reese G, et al. Salinas: A Scalable Software for High-performance Structural and Solid Mechanics Simulation [C] // Proc of the ACM/IEEE SC'02, 2002: 35.
- [23] Phillips J, Zheng G, Kumar S, et al. NAMD: Biomolecular Simulation on Thousands of Processors [C] // Proc of the IEEE/ACM SC'02, 2002: 36.
- [24] Akcelik V, Bielak J, Biros G, et al. High Resolution Forward and Inverse Earthquake Modeling on Terascale Computers [C] // Proc of the ACM/IEEE SC'03, 2003: 52.
- [25] Makino U, Daisaka H, Kokubo E, et al. Performance Evaluation and Tuning of GRAPE-6 Towards 40 'Real' Tflop/s [C] // Proc of the ACM/IEEE SC'03, 2003.
- [26] Adams M F, Bayraktar H H, Keaveny T M, et al. Ultras-

- calable Implicit Finite Element Analyses in Solid Mechanics with over a Half a Billion Degrees of Freedom[C]//Proc of the ACM/IEEE SC'04, 2004:34.
- [27] Vranas P, Bhanot G, Blumrich M, et al. The BlueGene/L Supercomputer and Quantum Chromodynamics[C]//Proc of the ACM/IEEE SC'06, 2006:50.
- [28] Wang L W, Lee B, Shan H, et al. Linear Scaling Divide-and-Conquer Electronic Structure Calculations for Thousand Atom Nanostructures[C]//Proc of the ACM/IEEE SC'08, 2008.
- [29] Ananthanarayanan R, Esser S K, Simon H D, et al. The Cat is Out of the Bag: Cortical Simulations with 109 Neurons, 1013 Synapses[C]//Proc of the ACM/IEEE SC'09, 2009:63.
- [30] Shaw D E, Dror R O, Salmon J K, et al. Millisecond-Scale Molecular Dynamics Simulations on Anton [C] // Proc of the ACM/IEEE SC'09, 2009:39.
- [31] Kozhevnikov A V, Eguiluz A G, Schulthess T C. Towards First-Principles Electronic Structure Simulations of Excited States and Strong Correlations in Nano- and Materials Science[C]//Proc of the ACM/IEEE SC'10, 2010.
- [32] Shimokawabe T, Aoki T, Takaki T, et al. Peta-Scale Phase-Field Simulation for Dendritic Solidification on the TSUB-AME 2.0 Supercomputer[C]//Proc of the ACM/IEEE SC'11, 2011.
- [33] Luisier M, Boykin T B, Klimeck G, et al. Atomistic Nanoelectronic Device Engineering with Sustained Performances Up to 1.44 PFlop/s[C]//Proc of the ACM/IEEE SC'11, 2011.
- [34] Bernaschi M, Bisson M, Endo T, et al. Petaflop Biofluidics Simulations on a Two Million-core System[C]//Proc of the ACM/IEEE SC'11, 2011.
- [35] Grinberg L, Morozov V, Fedosov D, et al. A New Computational Paradigm in Multiscale Simulations: Application to Brain Blood Flow[C]//Proc of the ACM/IEEE SC'11, 2011.
- [36] Kawai A, Fukushige T, Makino J. \$7.0/MFlops Astrophysical N-Body Simulation with Treecode on GRAPE-5[C] //Proc of the 1999 ACM/IEEE Conference on Supercomputing, 1999:67.
- [37] Kawai A, Fukushige T. \$158/GFlops Astrophysical N-Body Simulation with Reconfigurable Add-in Card and Hierarchical Tree Algorithm[C]//Proc of the ACM/IEEE SC'06, 2006:48.
- [38] Hamada T, Yokota R, Nitadori K, et al. 42 TFlops Hierarchical N-body Simulations on GPUs with Applications in Both Astrophysics and Turbulence[C]//Proc of the ACM/IEEE SC'09, 2009:62.
- [39] Hamada T, Nitadori K. 190 TFlops Astrophysical N-body Simulation on Cluster of GPUs[C] // Proc of the ACM/IEEE SC'10, 2010.
- [40] 张柏春. 科技革命及其对国家现代化的推动刍议 [J]. 科学与社会, 2012, 2(1):22-32.
- [41] TOP500 Supercomputing Sites[EB/OL]. [2012-05-12]. <http://www.top500.org>.



张理论(1975 -),男,河南邓州人,博士,副研究员,研究方向为大规模数值并行计算。**E-mail:**zll0434@sina.com

ZHANG Li-lun, born in 1975, PhD, associate research fellow, his research interest includes massively numerical parallel computing.



邓小刚(1960 -),男,四川绵阳人,博士,研究员,研究方向为计算空气动力学。**E-mail:**xgdeng2000@vip.sina.com

DENG Xiao-gang, born in 1960, PhD, research fellow, his research interest includes computational aerodynamics.