

加大对计算科学支持力度，应对艾级超算应用挑战

张云泉¹ 袁国兴²

¹(中国科学院软件研究所 北京 100190)

²(北京应用物理与计算数学所 北京 100088)

1 引言

以超级计算为工具的计算科学，已经渗透到科学研究与工程设计的各个层面，成为一个具有战略意义的新兴学科。计算科学已成为理论和实验之外，人类进行科学探索和工程研究的第三种手段。为此，美国总统信息技术顾问委员会PITAC在2005年6月向总统布什提交的题为《计算科学：确保美国的竞争优势》^[3]的报告中，建议美国政府制定长期计划，对计算科学进行长期资助，确保美国的竞争优势和国家安全。其所指的计算科学就是利用先进计算能力理解和求证复杂问题的学科，其已经成为科学研究领先、经济竞争优势和国家安全的至关重要因素。该委员会认为计算科学是21世纪最为重要的技术领域之一，因为它对整个社会的进步是不可或缺的。

我国的超级计算机研制近几年奋起直追，2009年，成为继美国之后，世界上第二个能够独立研制千万亿次峰值超级计算机的国家，超级计算机的研制水平日益提高，与国际水平的差距日益缩小。这得益于国家中长期科技计划的指导和国家863等科技计划的大力支持，也得益于国产厂商和科研人员的刻苦攻关和不懈努力与探索。虽然通过国家863项目紧密跟踪国际进展和趋势，积极部署相关的研制课题，千万亿次国产超级计算机在2009年底提前研制成功，比预期的时间提前了1年，历史性取得了世界第五名的位置^[1]；2010年底，部署的天河1A超级计算机，更是一举夺得世界第一的位置^[1]，取得历史性的突破。但是，由于在超级计算系统研发课题的部署中，不重视计算科学的发展，课题研究内容和研究经费配置不合理，对机器的研制周期缺乏科学论证，重视硬件系统的投入，忽视系统软件和应用软件的经费和人才投入，缺

乏有效的关键技术前期预研和积累机制，使得所研制出的超级计算机的应用和推广，由于缺乏成熟的系统软件和应用软件的支撑，成为一个非常具有挑战性的课题。在国产超级计算应用软件与国产超级计算机系统的发展水平差距越来越大的背景下，我国很有必要借鉴美国等西方发达国家发展计算科学和超级计算应用的成功经验，探索具有我国特色的计算科学学科发展与超级计算系统研制及应用推广均衡发展之路。

美国超级计算系统研制与应用的成功经验

美国的超级计算和计算科学的发展以能源部（department of energy, DOE）和DARPA为主，DOE为此还成立了专门的科学办公室。美国从1983年就制定了“战略计算机计划（strategy computer program, SCP）”；1993年实施了“高性能计算与通信（high-performance computing and communications, HPCC）计划”；1996年实施“加速战略计算创新（accelerated strategic computing innovation, ASCI）计划”及后续的“先进模拟和计算（advanced simulation and computing, ASC）计划”^[4]；2001年启动“基于先进计算的科学发现（found based on advanced computing science, SCIDAC）计划”^[8]专门用于发展超级计算所需要的物理模型、并行算法和并行软件；2006年启动“高生产率计算系统（high productivity computing systems, HPCS）计划”提高超级计算的并行软件生产率^[7]；2007年，美国NSF发布“通过千万亿次量级的仿真和分析加快科学工程领域的发现（PetaApps）计划”；2008年，美国能源部成立“先进科学计算研究办公室（advanced scientific computing research office, ASCR）”，专门负责统筹计算科学的发展，年度预算超过3.5亿美元；2009年启动“国际Exascale软件项目（international exascale

张云泉，研究员，博士生导师，主要研究领域为并行算法与并行软件。zyq@mail.rdeps.ac.cn。袁国兴，研究员。主要研究领域为大规模科学与工程计算的方法与软件。

software project, IESP) ”^[5], 以应对百万万亿超级计算机的软件挑战。

历史经验表明, 为了取得2008年在RoadRunner千万亿次超级计算机上的历史性突破, 美国在千万亿次系统的应用尚不明朗的情况下, 1994年就提前进行了部署和研讨, 启动了一系列科研计划, 瞄准这一目标提前开展研究工作, 花费了接近15年的时间, 才确保这一目标的如期实现。如今, 在千万亿次突破之后, 世界各国的目光无不转向超级计算机速度的下一个皇冠-Exascale (艾可萨次, 百万万亿, 10^{18}) 超级计算机。根据2009年美国的研究路线图, 美国大约在2015年开始艾可萨次超级计算机的部署, 并希望从2008年开始用比千万亿次的周期一半的时间 (也就是7年), 实现艾级超级计算机。

总结美国在超级计算研发上的成功经验, 其要点如下:

(1) 从国家层面协调超级计算系统软硬件的研制, 成立专门的领导机构和专家委员会, 通过面向全国的全面和深入调研, 拟定未来10到20年的发展规划和蓝图, 给以经费支持, 在政府指导下组织实施;

(2) 超级计算系统的规划、研制和部署形成流水线, 保证对超级计算研制的稳定支持, 缩短研制周期, 确保超级计算系统研制与应用长远发展;

(3) 在超级计算系统的研制经费分配中, 尊重计算科学的学科发展需要和规律, 非常重视对软件和硬件经费投入的平衡, 甚至更加重视软件的投资。软件的生命周期可以长达几十年, 而硬件的生命周期只有5年左右;

(4) 重视对超级计算系统软硬件研制队伍的建设。一般委托三到五家厂商通过多个阶段的竞争和逐级淘汰, 进行超级计算系统硬件的研制, 确保选出最优的方案和技术; 而超级计算系统的应用软件则依托实力雄厚的国家实验室或国家超级计算中心进行长期研制。这些国家实验室的人才队伍一般由计算机科学, 计算数学和应用领域的专家等符合计算科学学科交叉要求的人才组成;

(5) 超级计算系统的部署不强调一步到位, 而是在其生命周期中, 分阶段实施, 以确保其采用最新的技术和不断满足增长的应用需求, 以节省投资成本, 延长机器生命周期。

2 我国超级计算系统研制与应用的现状与存在的问题

上海超级计算中心是目前我国最早设立的超级计算中心之一, 峰值达到200万亿次规模的“魔方”系统(曙光5000A)在2009年5月正式投入运行。但该中心当前大部分软件作业难以有效利用几千以上的CPU, 从并行算法到并行软件都有很多工作要做。而中科院超级计算中心虽然通过千核应用课题支持了一批课题进行几千核甚至万核并行算法与软件研制的重点突破, 但目前运行的大部分作业的并行规模还是停留在成百上千个处理器甚至更低的量级。无法切实有效的发挥出百万亿次超级计算机的上万核并行的能力。即使是能够有效利用几百个处理器的并行软件, 大部分也是花费巨资购买的国外的商业应用软件。

因此, 要解决目前千万亿量级超级计算系统应用与推广的挑战性问题, 仅仅依靠研制千万亿量级的计算机硬件系统是不够的, 如果没有能够在千万亿量级超级计算机上有效运行的大规模并行的应用软件, 机器将很难发挥出应有的应用效果。

但我国长期以来在高性能计算领域重视硬件忽视算法和应用软件的研究。一方面在高性能计算机硬件的研制水平上与国外不断缩短差距, 另一方面由于长期的投入不足, 我国在高性能计算的算法和应用软件研究方面时断时续, 研究开发队伍的规模和水平一直无法得到长足的发展, 与国际水平相比差距越来越大。我国的算法和应用软件的研究水平已经开始严重制约我国高性能计算应用的深入和普及, 用户往往购买不起国外价格昂贵的应用软件, 甚至有些情况下, 有钱也购买不到能够满足应用需求的软件。国外的产品价格昂贵, 且形成垄断。国内的科研单位和公司存在各自为战, 小作坊式的开发的现象, 偏重于关键技术的突破, 而不重视产业化发展。更为严重的是, 目前国家还没有统一的国家机构和组织制定高性能计算的中长期发展规划, 进行统一的协调和组织。而美国的高性能应用软件的的基础研究和关键技术的研究开发方面不断制定目标明确的中长期研究规划和路线图, 对高性能计算的研发进行长期稳定的资助。其研发力量主要集中在几个国家实验室, 这些国家实验室的研发队伍配置齐全, 有稳定的支持, 人数众多, 专业分工明确。

总体来讲, 我国超级计算应用研究与开发明显滞后于高端计算机的发展已经是不争的事实, 应用的并行度普遍在成百上千量级, 应用到更大规模(万量级)的很少(并非没有需求)。超级计算应用软件的开

发和高效并行算法的研究尚不能与超级计算机的发展同步，在一定程度上存在为计算机“配”软件的思想。对应用的投入远远不够，应用研发的力量薄弱且分散，缺乏跨学科的综合型人才，从事超级计算应用软件研发的单位很少，企业界基本未介入，没有良好的、相互交流的组织渠道，缺乏统一的领导和协调机构等。

随着2010年底天河1A问鼎世界第一，以及若干台千万亿次超级计算机陆续投入使用，这一问题也显得更加严重和迫切。更为迫切的是，国家还计划在在十二五期间研制的十亿亿次超级计算机，很快也将提上议事日程。

3 相关建议

我国必须从国家层面统筹规划，成立统一的超级计算领导机构和专家委员会，重视计算科学学科和人才队伍建设，成立国家级的重点实验室或工程中心，加大对超级计算并行应用软件研发的投资力度，提前布局，协调发展，为下一步研制百万亿量级的高性能计算机做好应用相关的理论、关键技术和研究队伍的储备，这样才能确保所研制出来的超级计算机发挥出真正的战略性的作用。特提出如下建议：

(1) 重视计算科学的学科建设，有条件的大学可以成立计算科学系；在大学开设计算科学的课程，从大学生阶段就重视计算科学人才的培养，以做好人才队伍的建设和储备；

(2) 警惕近期全国各地超级计算中心纷纷上马的过热势头，避免重复投资和盲目投资，凡是要上超算中心的省份，要进行科学论证，确有必要的，可以加大支持力度。要全国一盘棋，统筹安排国家级超算中心的布局，做到既要符合当地当前的需要，又要对当地和周边区域的科技创新有一定的引领作用；

(3) 尽快筹建直接向国务院负责的国家级的超级计算战略领导小组，制订我国超级计算发展的中长期发展规划，统筹协调全国的超级计算软硬件的研制计划和经费安排，集中力量，避免人力和经费等资源的重复投资和浪费；

(4) 尽快成立超级计算领导小组领导下的高性能计算应用软件专家咨询委员会，统筹指导和规划国产高性能计算应用软件的中长期发展规划，制订切实可行的国产高性能应用软件发展计划；

(5) 在国家科技部等支持的超级计算系统的研

制课题中，要注意对硬件和软件的均衡投入和支持，硬件和软件的经费比例，要尽量做到一比一；

(6) 成立若干超级计算应用软件国家重点实验室或国家并行软件工程中心，组建学科全面的计算科学研究开发队伍，给以稳定的经费和课题支持，开展千万亿量级应用相关的基础问题和关键共性技术的研究，以支撑超大规模并行应用在数万处理器以上的计算机系统上能够算得出、算得快、算得可信。

参 考 文 献

- [1] Meuer H, Simon H, Strohmaier E, et al. TOP500 supercomputer sites [EB/OL]. <http://www.top500.org>
- [2] 张云泉, 孙家昶, 袁国兴, 等. 中国高性能计算机TOP100排行榜 [EB/OL]. <http://www.samss.org.cn>
- [3] Reed D A, et al. Computational science: ensuring america's competitiveness [EB/OL]. USA:PITAC, [2005-06-09]. http://www.nitrd.gov/pitac/reports/20050609_computational/computational.pdf
- [4] ASCI project [EB/OL].<http://www.lanl.gov/asc/>
- [5] Dongarra, J, Beckman P, et al. The international exascale software roadmap [EB/OL]. International Journal of High Performance Computer Applications, ISSN 1094-3420, 2011, 25(1). http://www.exascale.org/iesp/Main_Page
- [6] Zhang Y Q, Sun J C, Yuan G X, et al. Perspectives of china's HPC system development: a view from the 2009 china HPC TOP100 list [J]. Frontiers of Computer Science in China, 2010: 437-444.
- [7] DARPA high productivity computing systems program(HPCS) program [EB/OL]. http://en.wikipedia.org/wiki/High_Productivity_Computing_Systems
- [8] SciDAC project [EB/OL]. <http://www.scidac.gov/>