

# 上海交通大学 高性能计算中心年度报告 2015



## 主任寄语



《高性能计算中心年度报告 2015》和大家见面了，这是我们 2015 年工作的总结，展现了  $\pi$  超级计算机运行情况和用户使用  $\pi$  所取得的成绩。 $\pi$  团队和广大用户通过紧密合作、互相理解、互相支持，使得高性能计算服务更好地助力学校的高水平科研，体现了校级高性能计算平台在资源共享、稳定运维和精准服务等方面的巨大优势。

$\pi$  已经陪伴我们走过了第三个年头，高性能计算中心的服务也进入了需求高速增长期。与去年相比，2015 年 CPU 机时数增长了 7%，年利用率达到了 65%；GPU 机时数增长了 75%，年利用率达到了 58%，设备利用率处于国内外同行较好水平。

我们的团队也在实践中不断成长，推出了高附加值的服务。我们和物理与天文系、密西根学院等多个课题组进行了跨学科合作，协助用户针对  $\pi$  超级计算机进行代码调优；参与了学校转化医学、材料基因组等基础设施和研究项目的规划和建设工作；承担了国家 863 高性能计算相关的研究工作，成为中国超算创新联盟的理事成员，并与 Princeton 大学一起成功申请了美国 NSF 的国际跨学科合作 SAVI 项目。自身的技术能力不断提升，赢得了广大用户的信任和支持。

展望未来，我们的工作还有很大的改进空间。我们需要进一步增强设备运行的稳定性和可靠性，提升服务能力和水准，推动校级高性能计算平台计算能力和存储能力的升级换代。我相信有了学校的支持和广大用户的信任，我校高性能计算中心一定能在探索国内高校超算中心发展新模式的道路上继续前行，努力实现综合创新、人才培养和应用推广三个基地的建设目标。

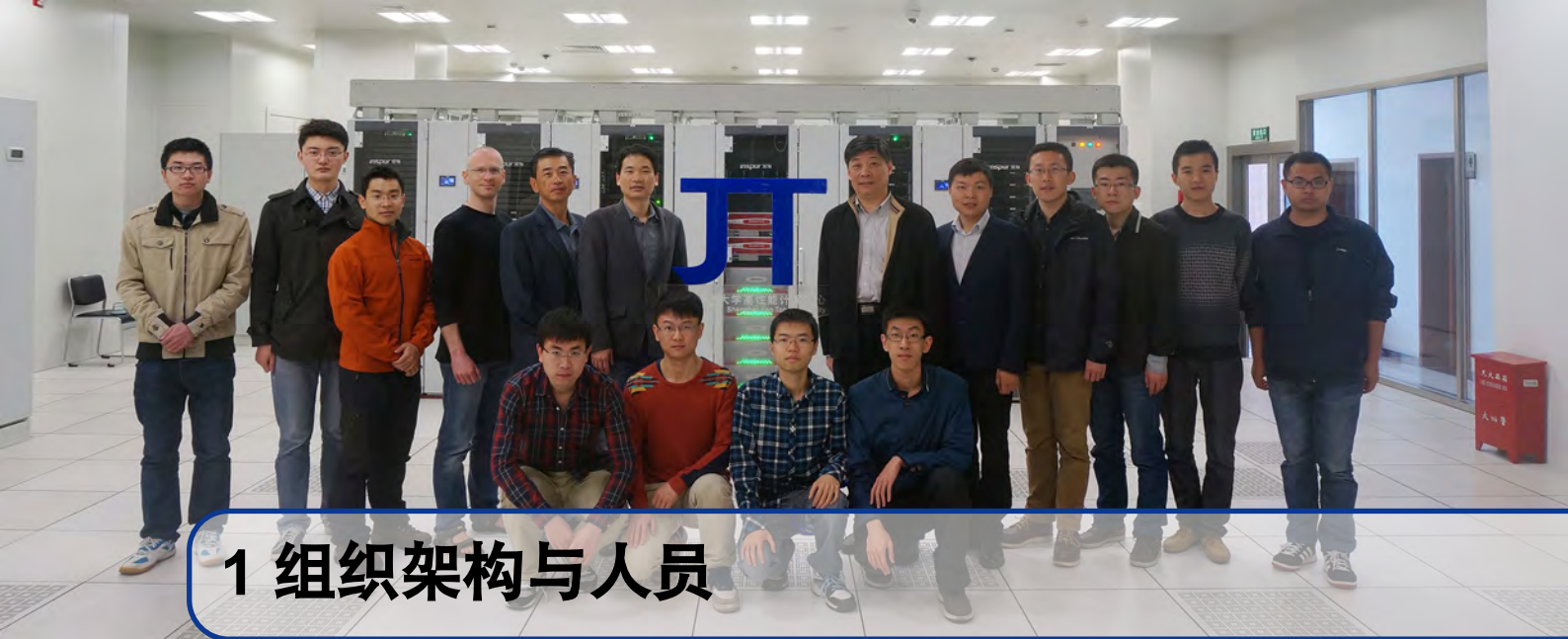
上海交通大学高性能计算中心主任  
二〇一六年四月十五日



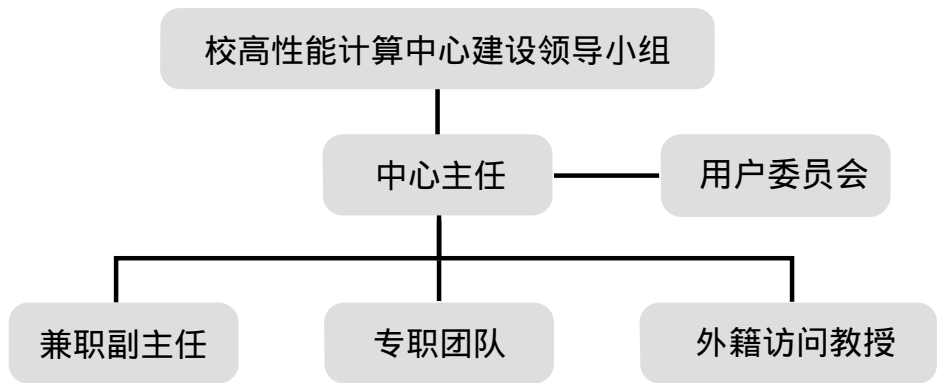


1	组织架构与人员	1
2	大事记	3
3	运维与资源分配	5
4	用户与成果统计	13
5	代码优化与研究	29
6	人才培养与培训	35
7	国际交流与合作	39
8	用户反馈与建议	43





# 1 组织架构与人员



**中心主任**全面负责中心的各项工作，由网络信息中心主任顾一众兼任。**专职团队**负责  $\pi$  集群的运维、用户支持和培训、以及用户代码调优等工作，共 5 名员工，其中 1 人为法籍：

林新华	副主任	在职博士	主持日常事务，负责国际化和研究
文敏华	计算专员	交大硕士	协助研究，负责学生竞赛，参与用户代码调优
韦建文	计算专员	交大硕士	负责系统运维，参与用户代码调优
王一超	计算专员	交大硕士	协助系统运维，负责用户培训，参与代码调优
Eric	计算专员	EPFL 博士	负责支持物理、材料领域等计算问题

**兼职副主任**负责协助和指导中心的服务及发展，共 4 位：

---

刘 洪	航空航天学院	教授，副院长
管海兵	网络信息中心	教授，书记、副主任
蔡申瓯	自然科学研究院	教授，院长
金先龙	机动学院	教授

---

**用户委员会**负责评审用户的机时申请，主任为蔡伟教授，共 7 位：

---

蔡 伟	自然科学研究院	教授
盛政明	物理与天文系	教授
魏冬青	生命科学技术学院	教授
金朝晖	材料学院	教授
刘兴钊	电信学院	教授
陆 皓	材料学院	教授
林新华	高性能计算中心	副主任

---

**外籍访问教授**负责协助中心提升科研实力和国际影响力，共 2 位：

---

Prof. William Tang	美国 Princeton 大学教授、美籍
Dr. Simon See	NVIDIA 亚太区技术总监，新加坡籍

---



## 1月

13日，中心首次深入学院开展 HPC 培训活动，在密西根学院由中心员工韦建文主讲“Getting Performance Boost Without Code Modification”专题讲座。

## 2月

25日，在  $\pi$  集群上进行计算的激光尾波加速研究成功发表于顶级期刊 Physical Review Letter。

## 3月

26日，美国 Argonne 国家实验室计算机科学家、MPI 标准制定的核心人员 Pavan Balaji 博士应邀为交大师生讲授 MPI 编程课程。

## 5月

5日，《高性能计算中心 2014 年度报告》正式发布，这是中心成立以来的第一份年度报告。

9-10日，中心与 Intel 联合举办了第四届 Intel-SJTU 华东地区 HPC 研讨会。华东地区十多所高校及企业近四十余位专家参加，促进了高性能计算的区域性交流与合作。

18日，上海交通大学 GPU 卓越中心 (GCOE) 与 NVIDIA 联合举办“GPU 加速深度学习研讨会”，吸引了超过 200 人参加本次研讨会。

17-25日，由中心员工文敏华带队的交大本科生代表队，赴山西太原参加 ASC15 世界大学生超算竞赛决赛，获得应用创新奖。



## 6 月

3 日, 澳大利亚国立大学计算机研究院主任 Alistair Rendell 教授一行访问交大, 与  $\pi$  团队进行了并行编程相关的学术交流。

6 月 26-7 月 3 日, 由于空调制冷系统以及存储控制器故障, 集群停机一周。

31 日, 上半年机时折扣申请截止。共收到 11 份有效申请, 分别来自物理与天文系、自然科学研究院、生命科学技术学院等 8 个院系。

## 7 月

10 日, 上海交通大学领导的国际团队在《细胞》和《柳叶刀》联合支持的开源杂志《EBioMedicine》发表关于“肠道菌群是儿童遗传性肥胖营养治疗的有效靶点”的重大发现论文。该项工作的高通量元基因组数据计算是在  $\pi$  集群上完成。

23 日, 中心举办第一次  $\pi$  集群用户基础培训。有 20 名集群新用户报名参加, 由中心员工王一超负责讲解并现场指导上机练习。

## 10 月

21 日, 发布修订版的高性能计算中心管理办法。本次修订主要改进了费用分担标准, 旨在使更多 GPU 用户享受到折扣。

30 日, 集群图形化节点上线测试, 为用户提供远程可视化服务。

## 11 月

9-12 日, 赴无锡参加 HPC China 2015。中心员工和所指导的学生共投中 4 篇论文, 数量位于全国高校前列。中心指导的两支学生队伍参加了由计算机学会高性能专业委员会和 Intel 联合主办的 PAC15 大赛, 分别获得性能优化组和最佳应用组的铜奖。

16-19 日, 赴美国 Austin 参加 SC15 大会, 作为大陆唯一参加布展的高校, 提升交大在国际高性能计算领域的影响力。

27-28 日, 中心组织的第二届  $\pi$  用户大会顺利举行。15 位  $\pi$  集群用户代表以及多位国内外高性能计算领域专家济济一堂, 展示自己在  $\pi$  上所做的工作, 并为如何进一步推动交大高性能计算发展献计献策。

## 12 月

1 日, 第四届 AMD 高性能计算奖学金申报开始。奖学金面向校内在  $\pi$  上进行高性能计算研究和应用的硕士生和博士生。奖励 5 名学生, 每人 1 万元。

31 日, 下半年机时折扣申请截止。共收到 7 份有效申请, 分别来自电信学院、船建学院、物理与天文系、医学院等 6 个院系。

## 3 运维与资源分配

### 3.1 $\pi$ 集群概述

$\pi$  的峰值计算能力为 336TFlops，其中 65% 的计算能力来自 GPU。2013 年 6 月建成时世界排名为 158，目前为国内高校第 4，上海地区第 2。取名为  $\pi$  有 4 层含义：

1.  $\pi$  在希腊文里有并行的涵义；
2.  $\pi$  是无限不循环的无理数，是人类理解无限的开始；
3.  $\pi$  的精确计算是人类使用计算机解决科学问题的代表；
4.  $\pi$  还可以看做为交通大学首字母 J 和 T 的组合。

$\pi$  的配置如下表所示。为了满足不同应用的需要，并在建成后陆续增加新的计算资源，目前的计算结点有 5 种不同类型：

节点类型	节点数目	节点配置	峰值性能
CPU 节点	332	2 颗 Intel Xeon E5-2670 CPU，64GB 内存	110TFLOPS
GPU K20 节点	50	2 片 NVIDIA K20 GPU 加速卡，64GB 内存	134TFLOPS
GPU K40 节点	5	2 片 NVIDIA K40 GPU 加速卡，64GB 内存	16TFLOPS
GPU K80 节点	15	2 片 NVIDIA K80 GPU 加速卡，96GB 内存	69TFLOPS
大内存节点	20	2 颗 Intel Xeon E5-2670 CPU，256GB 内存	7TFLOPS

## 3.2 公开运维信息

中心贯彻运维信息必须透明的原则，通过以下多种渠道及时公开资源利用和系统运行情况：

- 提供在线的实时利用率查询<sup>1</sup>，为国内唯一能提供该服务的高性能计算中心；
- 提供在线故障维修记录查询<sup>2</sup>，也是国内唯一；
- 建立  $\pi$  用户微信群，及时通告  $\pi$  的重要运维情况。

## 系统利用率

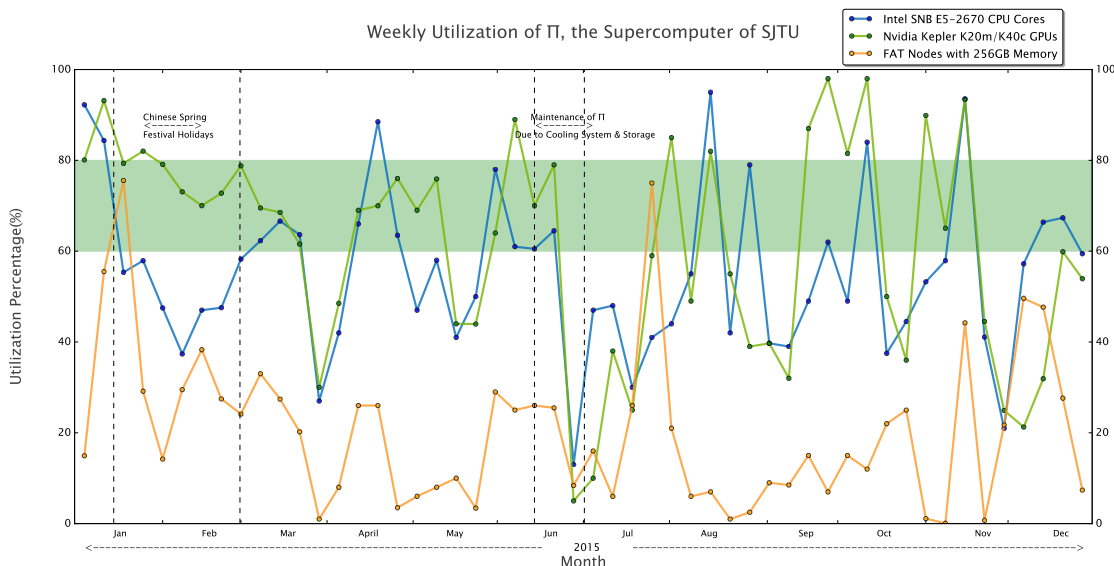


图 2015 年 1 月至 12 月的  $\pi$  集群系统利用率，其中蓝线表示 CPU 节点利用率，绿线表示 GPU 节点利用率，黄线表示大内存节点利用率。

- 健康的系统利用率应该在 60%-80% 之间（绿色阴影区域），过高则意味着系统过载，用户提交作业的等待时间变长。从上图可以看出，从整体上来看， $\pi$  集群的系统利用率在 2015 年都保持在较理想的区间。与 2014 年相比，2015 年提供的 CPU 年机时数增长了 7%，年平均利用率达到了 65%；GPU 年机时数增长了 75%，年平均利用率达到了 58%，设备利用率处于国内外同行较好水平。

<sup>1</sup><http://pi.sjtu.edu.cn/ganglia>

<sup>2</sup><http://pi.sjtu.edu.cn/log>

- 2015 年寒假期间机器利用率显著下降。主要是因为年关将近，缴费合同等手续来不及办理。新学期伊始，这些手续办完之后，利用率迅速回升到健康范围。
- 2015 年暑假因为大楼管道维护必需关停空调室外机，导致  $\pi$  集群下线，利用率骤降。
- GPU 结点的利用率与 CPU 结点几乎持平，有时候甚至还要高一些，这表明提供了 60% 左右计算能力的 GPU 能很好地支持生产型应用。
- 大内存的结点数量少（不到整机的 5%），仅供一些有大数据处理需求的用户使用（如生命科学技术学院），因此利用率波动会比较大。

## 计算量最大的项目、应用与院系

表 2015 年度  $\pi$  上 10 个计算量最大的项目

院系	PI	项目名称	实际使用机时
自然科学研究院	Jakob	Mechanism of membrane active antimicrobial peptides	240 万核小时 CPU 30 万卡小时 GPU
物理与天文系	盛政明	基于激光尾波场加速的可调谐超短 X 射线脉冲源的研究	280 万核小时 CPU
自然科学研究院	胡丹	肽-膜相互作用的分子动力学模拟和自由能计算	85 万核小时 CPU 10 万卡小时 GPU
材料学院	金朝晖	面向先进镁合金制造的多尺度模拟及高通量计算	150 万核小时 CPU
化学化工学院	孙淮	基于超级电容器的大容量储能体系及其应用	126 万核小时 CPU
物理与天文系	杨小虎	宇宙结构演化的理论研究	110 万核小时 CPU
物理与天文系	何峰	仄秒脉冲的产生	107 万核小时 CPU
数学系	金石	高超声速飞行器多尺度多物理运输问题的计算方法	95 万核小时 CPU
生命科学技术学院	陈海峰	microRNA 翻译上调的分子机制及实验验证	42 万核小时 CPU 5 万核小时大内存节点
生命科学技术学院	韦朝春	3000 个水稻基因组测序数据分析	40 万核小时 CPU 7 万核小时大内存节点

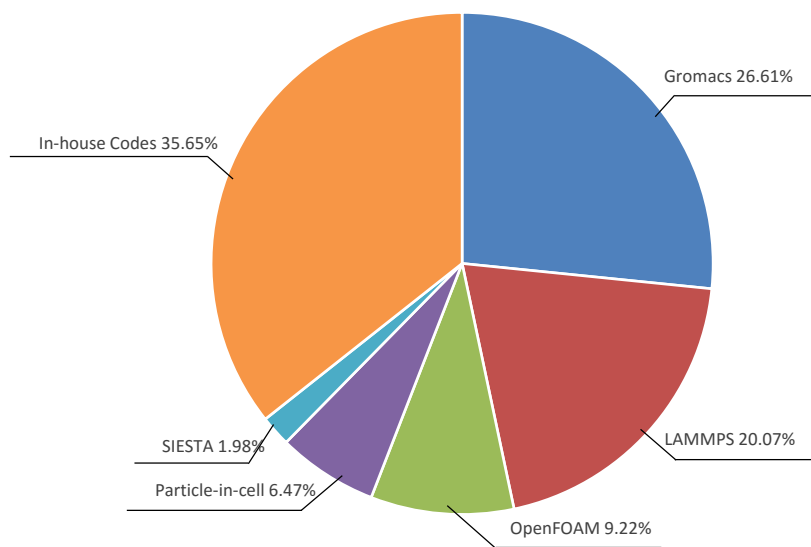


图 使用 CPU 最多的应用

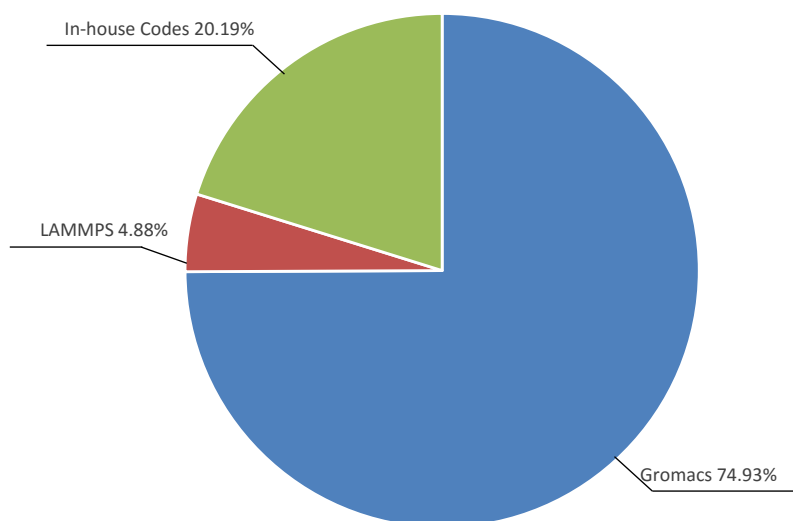


图 使用 GPU 最多的应用

### 3.3 公平分配资源

#### 分配原则

$\pi$  系统投入大, 运维成本高。为使其充分发挥效用, 在开放服务中, 我们遵循了以下两项原则:

1. 坚持**面向重点和广泛受益**原则。优先保证对大规模计算的需求, 特别是冲击国际前沿水平、涉及重大基础理论研究或涉及国民经济重大应用的国家级课题; 其次尽可能满足校内科研和教学中的中小规模计算需求; 少量考虑校外合作。
2. 坚持**费用分担和成果奖励**原则。通过学校补贴一部分、用户分担一部分的方式维持中心的可持续运行。对于重要的前沿探索和实质性合作研究项目以及取得重大成果的高水平课题, 经用户委员会评议可申请优惠。

#### 申请方式

为避免机时资源被滥用, 我们采用了费用分担的方法, 用户可以使用以下两种方式申请  $\pi$  的计算资源:

- 通过用户委员会申请。用户委员会负责分配  $\pi$  上约 70% 的资源。每年 7 月和 12 月接收用户申请, 对具有重大科学价值和研究前景的项目予以机时优惠支持。
- 直接向高性能计算中心申请为标准用户。中心可以分配约 30% 的资源。用户如果错过用户委员会的申请截止时间或是申请未被接受, 可以选择这种方式。

#### 机时奖励

2015 年, 根据成果奖励的原则, 我们公布了  $\pi$  用户机时奖励规章制度 (试行), 具体如下:

- 已缴纳经费的统计为 2014 年 1 月至 2015 年 4 月 (以年报定稿之日为准);
- 已发表论文数目统计以 2014 年年报为准, 且需注明该成果是在超级计算机  $\pi$  上运行的;
- 奖励标准为学校 A 类论文每篇奖励总使用机时的 8%, B 类奖励 4%, 其他论文归为 C 类奖励 2%。

表 2015 年  $\pi$  用户机时奖励结果统计

院系	PI	发表论文			奖励比例	使用机时 (万核小时)	奖励机时 (万核小时)
		A 类	B 类	C 类			
自然科学研究院	徐振礼	1	1	2	16%	5.0	0.8
物理与天文系	盛政明	1	3	2	24%	27.0	6.5
物理与天文系	何 峰	1	7	2	40%	3.3	1.3
物理与天文系	赵玉民		2		8%	4	0.3
物理与天文系	丁国辉		2		8%	1.8	0.14
数学系	金 石			1	2%	4	0.08
生命科学学院	赵一雷		1	2	8%	2.75	0.22
生命科学学院	陈海峰	1		5	18%	4.55	0.82
生命科学学院	韦朝春	1			8%	10.5	0.84
材料学院	金朝晖	1			8%	5	0.4
材料学院	孔令体			3	6%	5	0.3
航空航天学院	刘洪	3		3	30%	2.99	0.90
航空航天学院	李伟鹏		1	1	6%	2.5	0.15
船建学院	王本龙			1	2%	0.45	0.01
计算机系	俞凯			9	18%	2.67	0.48
合计		8	17	32		81.5	12.6

† 密西根学院以及自然科学研究院的个别用户由于另有合作协议，不在此次奖励名单中。

计算节点类型		校内标准用户	校内获得用户委员会奖励的用户	校外学术用户	非学术用户
CPU 节点 (64G 内存)	A: 每核小时 人民币 <b>0.1</b> 元	A*90%	A*60%	A*180%	A*360%
	包 50 万核小时 (含) 以上	A*80%	A*50%	A*160%	A*320%
GPU 节点	B: 每卡小时 人民币 <b>1.0</b> 元	B*100%	B*80%	B*200%	B*400%
	包 3 万 (含) 至 10 万卡小时	B*90%	B*60%	B*180%	B*360%
	包 10 万卡小时 (含) 以上	B*80%	B*50%	B*160%	B*320%
大内存节点 (256G 内存)	C: 每核小时 人民币 <b>0.2</b> 元	C*100%	C*80%	C*200%	C*400%
	包 10 万核小时 (含) 以上	C*80%	C*50%	C*160%	C*320%

图 2015 年 10 月，我们修订了费用分担标准，增大了使用 GPU 的折扣优惠

## 用户委员会批准支持的项目列表

编号	批准时间	院系	PI	项目名称	批准机时
1	2015 年 7 月	自然科学研究院	Jakob	Assembly and pore formation of maculatin	300 万核小时 CPU
2	2015 年 7 月	生命科学技术学院	魏冬青	质子跨膜机制的研究	100 万核小时 CPU 10 万卡小时 GPU
3	2015 年 7 月	航空航天学院	刘 洪	高超声速飞行器多尺度多物理运输问题的计算方法	90 万核小时 CPU
4	2015 年 7 月	化学化工学院	张万斌	新型手性配体的设计及金属不对称催化反应的开发	40 万核小时 CPU
5	2015 年 7 月	自然科学研究院	洪 亮	细胞色素蛋白中生物功能相关的大尺度关联运动随温度演变	45 万核小时 CPU
6	2015 年 7 月	物理与天文系	何 峰	阿秒激光物理和量子相干控制	100 万核小时 CPU
7	2015 年 7 月	物理与天文系	孙 弘	纳米结构对超硬材料强度激增效应新机制的第一性原理计算探索	150 万核小时 CPU
8	2015 年 7 月	材料学院	沈 耀	亚微米结构材料变形初期高硬化率机理的位错动力学模拟	100 万核小时 CPU
9	2015 年 7 月	医学院	陈赛娟	急性淋巴细胞白血病致病分子机制研究	4 万核小时 CPU
10	2015 年 12 月	海洋工程国家重点实验室	田新亮	港珠澳大桥管隧段的整体布放过程水动力仿真建模与计算	40 万核小时 CPU
11	2015 年 12 月	电子信息与电气工程学院	俞 凯	分布式大规模声学模型	1.5 万核小时 CPU 4 万卡小时 GPU
12	2015 年 12 月	环境科学与工程学院	程 真	长三角细颗粒物重污染过程中二次无机气溶胶的形成路径与来源解析研究	2 万核小时 CPU
13	2015 年 12 月	生命科学技术学院	陈海峰	天然无规蛋白质分子力场的开发与运用研究	10 万卡小时 GPU
14	2015 年 12 月	物理与天文系	於陆勒	激光等离子体相互作用中有关粒子加速、输运和辐射的大规模数值模拟研究	200 万核小时 CPU
15	2015 年 12 月	物理与天文系	罗卫东	拓扑绝缘体电子结构的磁性调控	200 万核小时 CPU
16	2015 年 12 月	医学院	陈赛娟	急性淋巴细胞白血病致病分子机制研究	100 万核小时 CPU





图 2015 年院系机时使用分布图

## 4 用户与成果统计

### 4.1 用户

截止至 2015 年底， $\pi$  上的标准用户总计 134 位，是 2014 年用户数的两倍多，覆盖了校内的 15 个院系。其中用户获得的荣誉称号如下（排名不分前后）：

荣誉称号	获得者	院系
两院院士 (1 人)	陈赛娟	医学院
千人计划 (1 人)	金 石	数学系
长江学者 (6 人)	金 石 盛政明 唐克轩 杨小虎 张大兵	数学系 物理与天文系 农业与生物学院 物理与天文系 生命科学技术学院
杰出青年 (5 人)	金 石 盛政明 杨小虎 周永丰 张大兵	数学系 物理与天文系 物理与天文系 化学化工学院 生命科学技术学院
青年千人 (6 人)	陈 民 Jakob 沈泳星 翁苏明 俞 凯 应文俊	物理与天文系 自然科学研究院 密西根联合学院 物理与天文系 电信学院 数学系
美国数学学会会士 (1 人)	金 石	数学系
美国物理学会会士 (1 人)	盛政明	物理与天文系
美国力学学会会士 (1 人)	陈谦斌	密西根联合学院

表 2015 年基于  $\pi$  完成的论文统计

院系	论文数		
	已发表	已提交或计划提交	小计
生命科学技术学院	8	5	13
物理与天文系	8	3	11
化学与化工学院	7	2	9
自然科学研究院	5	4	9
航空航天学院	4	2	6
医学院	1	0	1
合计	33	16	49

† 数据来自用户填写的反馈表。

## 4.2 论文

### 物理与天文系

- [1] Guo-Bo Zhang et al. "Acceleration and evolution of a hollow electron beam in wakefields driven by a Laguerre-Gaussian laser pulse". In: *Physics of Plasmas* (**IF=2.142**) 23.3 (2016), p. 033114.
- [2] Guo-Bo Zhang et al. "Acceleration of on-axis and ring-shaped electron beams in wakefields driven by Laguerre-Gaussian pulses". In: *Journal of Applied Physics* (**IF=2.183**) 119.10 (2016), p. 103101.
- [3] Min Chen et al. "Tunable synchrotron-like radiation from centimeter scale plasma channels". In: *Light: Science & Applications* 5.1 (2016), e16015.
- [4] M Mirzaie et al. "Demonstration of self-truncated ionization injection for GeV electron beams". In: *Scientific reports* (**IF=5.578**) 5 (2015).
- [5] Ming Zeng et al. "Multichromatic Narrow-Energy-Spread Electron Bunches from Laser-Wakefield Acceleration with Dual-Color Lasers". In: *Physical review letters* (**IF=7.512**) 114.8 (2015), p. 084801.
- [6] SM Weng et al. "Dense blocks of energetic ions driven by multi-petawatt lasers". In: *Scientific reports* (**IF=5.578**) 6 (2016).
- [7] KF Zhang et al. "Vectorial mapping of noncollinear antiferromagnetic structure of semiconducting FeSe surface with spin-polarized scanning tunneling microscopy". In: *Applied Physics Letters* (**IF=3.302**) 108.6 (2016), p. 061601.

- [8] Ping Li and Weidong Luo. “A new structure of two-dimensional allotropes of group V elements”. In: *arXiv preprint arXiv:1511.07173* (2015).

## 生命科学技术学院

- [9] Yan-Bing Qi et al. “Multicomponent kinetic analysis and theoretical studies on the phenolic intermediates in the oxidation of eugenol and isoeugenol catalyzed by laccase”. In: *Physical Chemistry Chemical Physics* (**IF=4.439**) 17.44 (2015), pp. 29597–29607.
- [10] Chenhong Zhang et al. “Dietary modulation of gut microbiota contributes to alleviation of both genetic and simple obesity in children”. In: *EBioMedicine* 2.8 (2015), pp. 966–982.
- [11] Zhiqiang Hu et al. “Revealing Missing Human Protein Isoforms Based on Ab Initio Prediction, RNA-seq and Proteomics”. In: *Scientific reports* (**IF=5.578**) 5 (2015).
- [12] Wei Ye et al. “Test and Evaluation of ff99IDPs Force Field for Intrinsically Disordered Proteins”. In: *Journal of chemical information and modeling* (**IF=3.738**) 55.5 (2015), pp. 1021–1029.
- [13] Dingjue Ji, Wei Ye, and HaiFeng Chen. “Revealing the binding mode between respiratory syncytial virus fusion protein and benzimidazole-based inhibitors”. In: *Molecular BioSystems* (**IF=3.21**) 11.7 (2015), pp. 1857–1866.
- [14] Lishi Xu et al. “Recognition Mechanism between Lac Repressor and DNA with Correlation Network Analysis”. In: *The Journal of Physical Chemistry B* (**IF=3.302**) 119.7 (2015), pp. 2844–2856.
- [15] Ke Wu et al. “Selectivity Mechanism of ATP-Competitive Inhibitors for PKB and PKA”. In: *Chemical biology & drug design* (**IF=2.485**) 86.1 (2015), pp. 9–18.
- [16] Jinmai Zhang et al. “Synergistic Modification Induced Specific Recognition between Histone and TRIM24 via Fluctuation Correlation Network Analysis”. In: *Scientific reports* (**IF=5.578**) 6 (2016).

## 自然科学研究院

- [17] Tao Du, Dan Hu, and David Cai. “Outflow Boundary Conditions for Blood Flow in Arterial Trees”. In: *PloS one* (**IF=3.8**) 10.5 (2015), e0128597.

- [18] Tao Du, Dan Hu, and David Cai. “Fast Numerical Method for Modeling One-dimensional Blood flow in Arterial Trees”. In: *Journal of Computational Physics (IF=2.434)* 314 (2016), pp. 450–464.
- [19] Yihao Liang, Zhenli Xu, and Xiangjun Xing. “A multi-scale Monte Carlo method for electrolytes”. In: *New Journal of Physics (IF=3.558)* 17.8 (2015), p. 083062.
- [20] Zecheng Gan et al. “Comparison of efficient techniques for the simulation of dielectric objects in electrolytes”. In: *Journal of Computational Physics (IF=2.434)* 291 (2015), pp. 317–333.
- [21] Sanjay K Upadhyay et al. “Insights from micro-second atomistic simulations of melittin in thin lipid bilayers”. In: *The Journal of membrane biology (IF=2.457)* 248.3 (2015), pp. 497–503.

## 航空航天学院

- [22] Bin Zhang and Jian-Hang Wang. “A short note on the counter-intuitive spurious behaviors in stiff reacting flow”. In: *Journal of Computational Physics (IF=2.434)* 291 (2015), pp. 52–59.
- [23] Bin Zhang, Hong Liu, and Shi Jin. “An asymptotic preserving Monte Carlo method for the multispecies Boltzmann equation”. In: *Journal of Computational Physics (IF=2.434)* 305 (2016), pp. 575–588.
- [24] Bin Zhang and Jian Hang Wang. “A Fraction Step Method Based on Adaptive Reaction Time Step and Its Application for High-speed Stiff Reacting Flows”. In: *Journal of Aeronautics, Astronautics and Aviation* 47.1 (2015), pp. 1–7.
- [25] Bin Zhang, Lin Ying Li, and Yong Xing Mi. “Rarefaction Effect on the Aerodynamic Performance of Scramjet Inlet in Hypersonic Flow”. In: *Journal of Aeronautics, Astronautics and Aviation* 47.4 (2015), pp. 357–362.

## 化学化工学院

- [26] Fenglei Cao and Huai Sun. “Transferability and Nonbond Functional Form of Coarse Grained Force Field—Tested on Linear Alkanes”. In: *Journal of chemical theory and computation (IF=5.498)* 11.10 (2015), pp. 4760–4769.

- [27] Zhifeng Jing, Liang Xin, and Huai Sun. "Replica exchange reactive molecular dynamics simulations of initial reactions in zeolite synthesis". In: *Physical Chemistry Chemical Physics* (**IF=4.493**) 17.38 (2015), pp. 25421–25428.
- [28] Zhe Shen and Huai Sun. "Prediction of Surface and Bulk Partition of Nonionic Surfactants Using Free Energy Calculations". In: *The Journal of Physical Chemistry B* (**IF=3.302**) 119.51 (2015), pp. 15623–15630.
- [29] Zhao Jin et al. "Hierarchical atom type definitions and extensible all-atom force fields". In: *Journal of computational chemistry* (**IF=3.589**) (2015).
- [30] Li-Tao Zhu et al. "Filtered Model for the Cold-Model Gas-Solid Flow in a Large-Scale MTO Fluidized Bed Reactor". In: *Chemical Engineering Science* (**IF=2.337**) (2016).
- [31] Mao Quan et al. "Pd (ii)-catalyzed asymmetric addition of arylboronic acids to cyclic N-sulfonyl ketimine esters and a DFT study of its mechanism". In: *Organic Chemistry Frontiers* 2.4 (2015), pp. 398–402.
- [32] Qun He et al. "Palladium-Catalyzed Asymmetric Addition of Arylboronic Acids to Nitrostyrenes". In: *Organic letters* (**IF=6.364**) 17.9 (2015), pp. 2250–2253.

## 医学院

- [33] Jianrong Xu et al. "Structural determinants for the interactions between muscarinic toxin 7 and muscarinic acetylcholine receptors". In: *Journal of Molecular Recognition* (**IF=2.151**) 28.4 (2015), pp. 239–252.

## 已提交或计划提交论文

- [34] Shijie Li, Xiaohu Yang, and et al. "An empirical model to form and evolve galaxies in dark matter halos". In: *RAA* (2016).
- [35] Y. Y. Cheng, Y. M. Zhao, and A. Arima. "Nucleon-pair states of semi-magic N=82 isotones". In: *Physical Review C* (**IF=3.733**) (2016).
- [36] Y. Y. Cheng et al. "Nucleon-pair states of semi-magic Sn isotopes based on realistic effective interaction". In: *Physical Review C* (**IF=3.733**) (2016).
- [37] G. et al. Wu. "Diminution of the gut resistome after a gut microbiota-targeted dietary intervention in obese children". In: *Scientific Report* (**IF=5.578**) (2016).

- [38] et. al Li X. "RNA-seq of embryonic cells and germ cells shows direct evidence for dosage compensation in mammals". In: *Nature Communication* (**IF=11.47**) (2016).
- [42] Dan Hu and David Cai. "Adaptation: the Optimization of Biological Transport Networks". In: *Science Sponsored Collection: Pushing the boundaries of scientific research: 120 years of addressing global issues* (2016).
- [43] E. Luijten Z. Gan S. Jiang and Z. Xu. "A hybrid method for systems of closely spaced dielectric spheres and ions". In: *SIAM J. Sci. Comp.* (**IF=1.854**) (2016).
- [44] S. Zhao M. Ma and Z. Xu. "Investigation of dielectric decrement and correlation effects on electric double-layer capacitance by self-consistent field model". In: *Communications in Computational Physics* (**IF=1.943**) (2016).
- [45] M. Ma P. Liu and Z. Xu. "Depletion induced like-charge attraction: Understanding from a self-consistent field model with dielectric effect". In: (2016).
- [46] Chen H and Zhang B. "On the non-Rankine-Hugoniot shock zone of the Mach reflection in hypersonic rarefied flows". In: *Journal of Spacecraft and Rockets.* (**IF=.533**) (2016), accepted.
- [47] Yu B et al. "An approach to obtain the correct shock speed for the Euler equations with stiff detonation". In: *Communications in Computational Physics* (**IF=1.943**) (2016).
- [48] Chunyang Yu and et al. "Molecular Dynamics Simulation Studies of Hyper-branched Polyglycerols and Their Encapsulation Behaviors of Small Drug Molecules". In: (2016).
- [49] Wanbin Zhang and et al. "Ni(II)-Catalyzed Asymmetric Addition of Arylboronic Acids to Cyclic Imines and a DFT Study of Its Mechanism." In: *J. Am. Chem. Soc.* (2016).

表 2015 年基于  $\pi$  正在进行的项目统计

院系	973	基金委			863	其他	小计
		重点	面上	青年			
物理与天文系	4	0	4	3	2	4	17
自然科学研究院	0	0	1	0	0	4	5
生命科学技术学院	0	1	1	0	1	1	4
医学院	0	0	0	1	0	1	2
化学化工学院	1	0	0	1	0	0	2
材料学院	0	0	2	0	0	0	2
航空航天学院	0	1	0	0	0	1	2
电信学院	0	0	1	1	0	0	2
环境科学与工程学院	0	0	0	0	0	1	1
密西根联合学院	0	0	0	1	0	0	1
合计	5	2	10	7	3	13	38

† 数据来自用户填写的反馈表。

### 4.3 项目

#### 物理与天文系

1. 张杰, 973 A 类项目课题四, “极端相对论光场驱动的粒子加速新机制探索”, 2013CBA01500, 2013–2018。
2. 向导, 青年 973 项目, “基于逆康普顿散射的高增益超快 x 光源若干前沿问题研究”, 2015CB859700, 2015–2019。
3. 罗卫东, 国家“青年千人”计划, “凝聚态物理及功能材料的第一性原理研究”, 2012–2016。
4. 罗卫东, 国家自然科学基金面上项目, “拓扑绝缘体电子结构磁性调控的第一性原理研究”, 2015–2018。
5. 赵玉民, 国家自然科学基金 No.11225524, No. 11505113, “单幻核配对态”, 2013–2017。
6. 赵玉民(子课题参加者), 973 项目 (NO.2013CB834401), “原子核质量的描述和预言”, 2013–2017。
7. 赵玉民, 国家自然科学基金杰出青年基金 (NO. 11225524), “原子核结构理论”, 11225524, 2013–2016。
8. 盛政明, 国家自然科学基金面上项目, “极端光场驱动超短相干强 X 射线源产生”, 11374210, 2014–2017。



9. 杨小虎, 973 项目, “ELUCID 系列数值模拟”。
10. 何峰, 国家自然科学基金面上项目, “阿秒激光物理和量子相干控制”。
11. 陈民, 国家自然科学基金面上项目, “基于激光尾波场加速的桌面伽玛射线辐射源”, 11374209, 2014–2017。
12. 陈民, 国家自然科学基金青年项目, “激光尾波场加速中横向光离化注入对电子的控制”, 11205101, 2013–2015。
13. 陈民, 上海市浦江人才项目, “基于激光尾波场加速的可调谐超短 X 射线脉冲源的研究”, 2013–2015。
14. 陈民, 863 项目, “激光等离子体相互作用及粒子加速与辐射”, 2012–2016。
15. 陈民, 863 项目, “利用激光等离子体相互作用产生相干高频辐射”, 2012–2015。
16. 於陆勒, 国家自然科学基金青年项目, “激光尾波场加速中的双色光离化电子注入”, 11405107, 2015–2017。
17. 翁苏明, 国家自然科学基金青年项目, “惯性约束核聚变的离子驱动快点火方案理论与数值模拟研究”, 11405108, 2015–2017。

## 自然科学研究院

18. Jakob, Young 1000 Talent grant, “Deciphering the atomic detail interactions of antimicrobial peptides with lipid bilayers”。
19. 胡丹, 国家自然科学基金, “膜-肽相互作用及自由能计算”, 11471213, 2015–2018。
20. 胡丹, 国家自然科学基金面上项目, “多步自由能算法及其在膜 - 肽相互作用中的应用”, 11471213, 2015–2018。
21. 徐振礼, 国家自然科学基金, “含多体效应的电荷输运方程的分析和计算”, No. 11571236, 2016–2019。
22. 徐振礼, 医工交叉研究, “带电软物质的计算和模拟”, 2015–2017。

## 生命科学技术学院

23. 赵立平, 国家自然科学基金重点项目, “Medical potential of gut microbiota in human health”, 31330005, 2014–2018。
24. 韦朝春, 863 项目 2014AA021502, “微生物组学数据集成及分析的关键技术研发”, 2014–2016。
25. 韦朝春, 国家自然科学基金面上项目 61472246, “元基因组中复杂结构的序列模

块寻找及其功能分析”，2015–2018。

26. 陈海峰，国家自然科学基金，“天然无规蛋白的环境特异性力场的发展与应用研究”。

## 医学院

27. 张健，万人计划“青年拔尖人才”项目，“The Structural Basis of Oncogenic Mutations G12, G13 and Q61 in Small GTPase K-Ras4B”。
28. 徐见容，国家自然科学基金科学青年基金项目，“M1 胆碱受体激动剂对  $G_{\alpha q}$ 、 $G_{\alpha s}$  及  $\beta$ -arrestin 信号通路的偏向性激活及其分子机制”，81503174。

## 化学化工学院

29. 罗正鸿，973 子课题，“多尺度层次下的循环硫化床中甲醇定向转制烯烃过程的化学工程研究”，2012CB215000。
30. 周永丰，国家自然科学基金青年基金，“超支化聚合物水溶液中构象行为和组装行为的计算机模拟研究”。

## 材料学院

31. 金朝晖，国家自然科学基金面上项目，“面向先进镁合金制造的多尺度模拟及高通量计算：强韧化机制”。
32. 沈耀，国家自然科学基金面上项目，“亚微米结构材料变形初期高硬化率机理的研究”，51471107，2015–2018。

## 航空航天学院

33. 张斌，国家自然科学基金重点项目，博士后基金，“高超声速飞行器多尺度多物理输运问题的计算方法”。
34. 李伟鹏 (Co-PI)， “Novel non-intrusive measurements and simulations of supersonic flow and noise”，新加坡教育部支持的国际合作课题 (MoE Tier 2, MOE2014-T2-1-002)，2015–2017。

## 电信学院

35. 俞凯, 国家自然科学基金委优秀青年科学基金, “基于深度学习的大规模声学模型训练及可视化分析”, 61222208, 2012–2015。
36. 顾海华, 国家自然科学基金青年基金, “密码学中的椭圆曲线理论研究”, 61202372, 2013–2015。

## 环境科学与工程学院

37. 程真, 国家科技支撑计划子课题, “长三角细颗粒物重污染过程中二次无机组分的形成路径与来源解析”。

## 密西根联合学院

38. 沈泳星, 国家自然科学基金青年科学基金, “基于万能网格及相场的水力压裂多尺度数值模拟方法”, 11402146, 2015–2017。

## 4.4 典型应用

### 应用一：大型客机缝翼噪声机理的大涡模拟研究

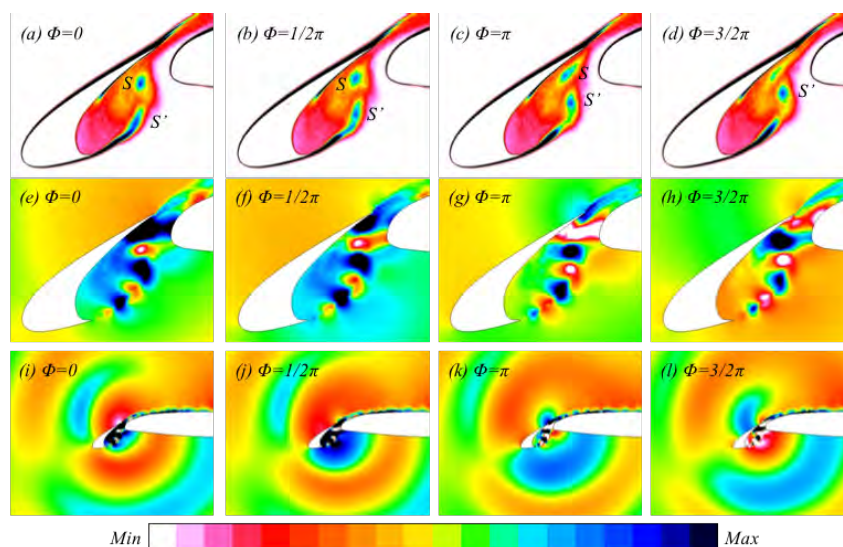


表 团队成员

李伟鹏	航空航天学院特别副研究员
刘洪	航空航天学院教授

气动噪声是大型客机研制过程中的关键问题之一，直接关系到大型客机的适航、环保性（声污染）和安全性（声疲劳）。利用上海交大的  $\pi$  系统，针对大型客机的缝翼噪声的产生机理，开展了大规模的数值计算工作，揭示了缝翼中的多尺度、非定常、复杂的流动现象，阐述了层/湍流边界层分离、剪切层混合和剪切层-壁面干扰等典型的流动问题，研究了涡声耦合、声源识别、声传播与声辐射等重要声学问题。利用高可靠性大涡模拟，实现缝翼流场/声场一体化研究，阐述流场涡系结构的生成与演化规律，掌握缝翼噪声的产生位置、强度和频谱特性。利用先进的数据挖掘方法，深度分析缝翼非定常流动与噪声之间的多尺度时空关联，揭示缝翼噪声的产生机理。相关的研究成果发表于 AIAA Journal 和 Journal of Aircraft 等高水平期刊上。

## 应用二：高通量计算模拟在材料基因组工程新材料预测中的应用

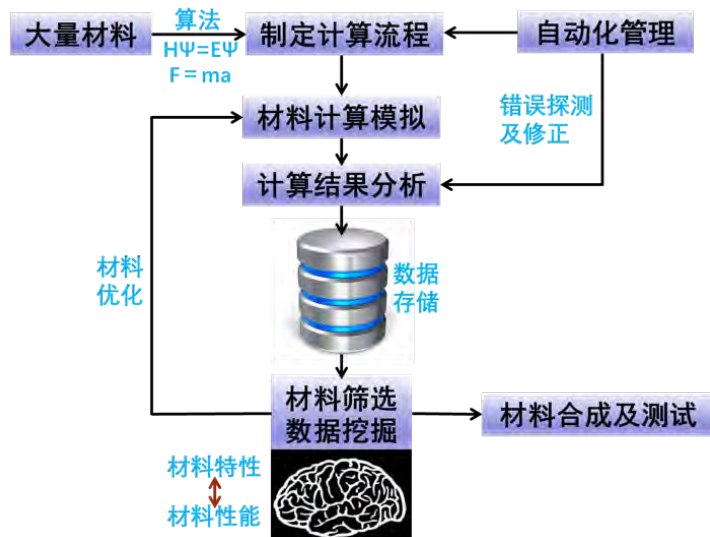


表 团队成员

张澜庭	材料科学与工程学院教授
孔令体	材料科学与工程学院副教授
朱虹	密西根联合学院助理教授
罗卫东	物理与天文系特别研究员
Éric Germaneau	高性能计算中心工程师
刘科	密西根联合学院博士研究生
邬嘉义	密西根联合学院博士研究生
戚文	密西根联合学院博士研究生

2011年，美国为复兴其制造业，正式启动“材料基因组计划”(Materials Genome Initiative)，旨在综合运用高通量计算、高通量实验和数据库技术，实现材料科学家“Materials by design”的梦想，并将新材料的开发周期和成本都降低一半。

高通量材料计算模拟作为该计划的重要组成部分已经在多个材料研究领域发挥了重要作用，2015年4月成立的上海交通大学材料基因工程联合研究中心，把组建Materials Project Asia Hub高通量计算平台作为三个基石方向之一，开展了在 $\pi$ 上搭建和调试高通量计算平台的工作。上海交通大学高通量计算团队来自材料学院、密西根学院、物理与天文系和高性能计算中心等多个院系，在以交大讲席教授、美国加州大学伯克利分校的Ceder教授为首Materials Project团队的指导下，以及高性能计算团队的帮助下，顺利完成了计算体系的构建，如今已开始镁合金表面、热电材料相关性计算的高通量计算模拟。

在2016年本项目将在热电材料的筛选扩大到1000个材料，并开展铁磁、结构材料相关的高通量计算流程设计，并进行新材料的理论预测，与此同时将加大对数据库的建设及展示。

## 应用三：基于深度神经网络的大规模声学模型训练

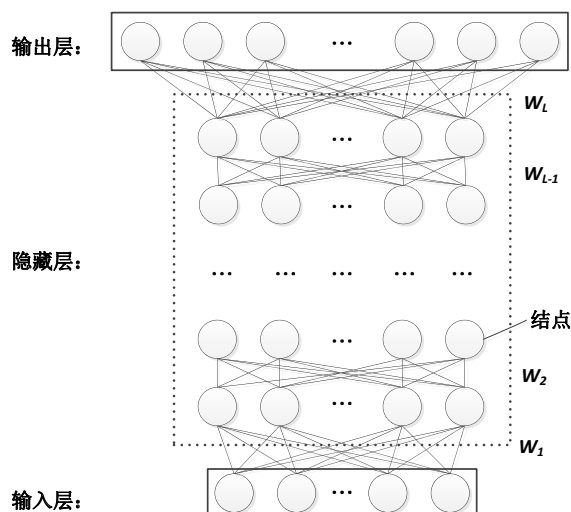


表 团队成员

俞 凯	计算机科学与工程系特别研究员
谭 天	计算机科学与工程系博士研究生
刘 奇	计算机科学与工程系博士研究生
项 煦	计算机科学与工程系硕士研究生
毕梦霄	计算机科学与工程系硕士研究生
陈瑞年	计算机科学与工程系本科生

语音是人和机器最便捷的交互方式之一。随着神经网络等最新理论的发展与应用, 语音交互技术中的语音识别已经达到了实用的阶段。谷歌, 微软, 苹果等公司近年来都相继推出了面向语音识别的商业产品, 并且都使用了基于神经网络的大规模声学模型。基于神经网络的声学模型, 如何对其有效训练是重点。一方面训练的数据相比十年前提高了 10 倍以上, 从数百小时的训练数据到现在上千小时的训练数据。另一方面, 深度神经网络的模型参数量巨大, 以一个七层的神经网络为例, 如果每层含有 2048 个隐层节点, 那么整体参数量为三千万。

本课题就各类深度神经网络的声学模型训练问题, 在上海交通大学的  $\pi$  高性能计算机上, 测试和搭建了数千小时语音数据的分布式大规模声学模型训练平台, 并对神经网络的大规模并行训练、聚类自适应训练、可视化分析进行了研究。并在语音识别的顶级刊物上发表数篇论文。另外我们还将进行神经网络训练工具包的开发, 涵盖各种神经网络, 包括卷积神经网络, 循环神经网络等。

## 应用四：3000 个水稻基因组测序数据比较分析



表 团队成员

---

张大兵	生命科学技术学院教授、长江学者、杰出青年
韦朝春	生命科学技术学院教授、浦江人才、教育部新世纪优秀人才
石建新	生命科学技术学院副教授
胡智强	生命科学技术学院博士研究生
孙 晨	生命科学技术学院硕士研究生

---

水稻是世界最重要的农作物和重要的模式生物之一。本项目收集了目前全世界 3024 个优良水稻品种，进行基因组测序。测序平均覆盖率近 15 倍，原始测序数据量为 17 TB。本项目是目前最全面、最广泛的水稻基因组研究。通过对这些水稻基因组比较分析，揭示水稻基因组的演变，探究影响水稻产量相关性状的分子机制，对水稻研究具有重要的理论指导意义和实际应用价值。

在  $\pi$  超级计算机上完成了基于 3000 水稻基因组测序数据的水稻泛基因组（即基因在不同水稻品系中的有无情况）分析（使用 CPU 核时超过 130 万，硬盘存储空间超过 100TB）。相关工作预计可以发表 2 篇以上生物学领域高水平学术论文。

### 应用五：道微生物组在人体代谢性疾病发生发展中作用的精细研究

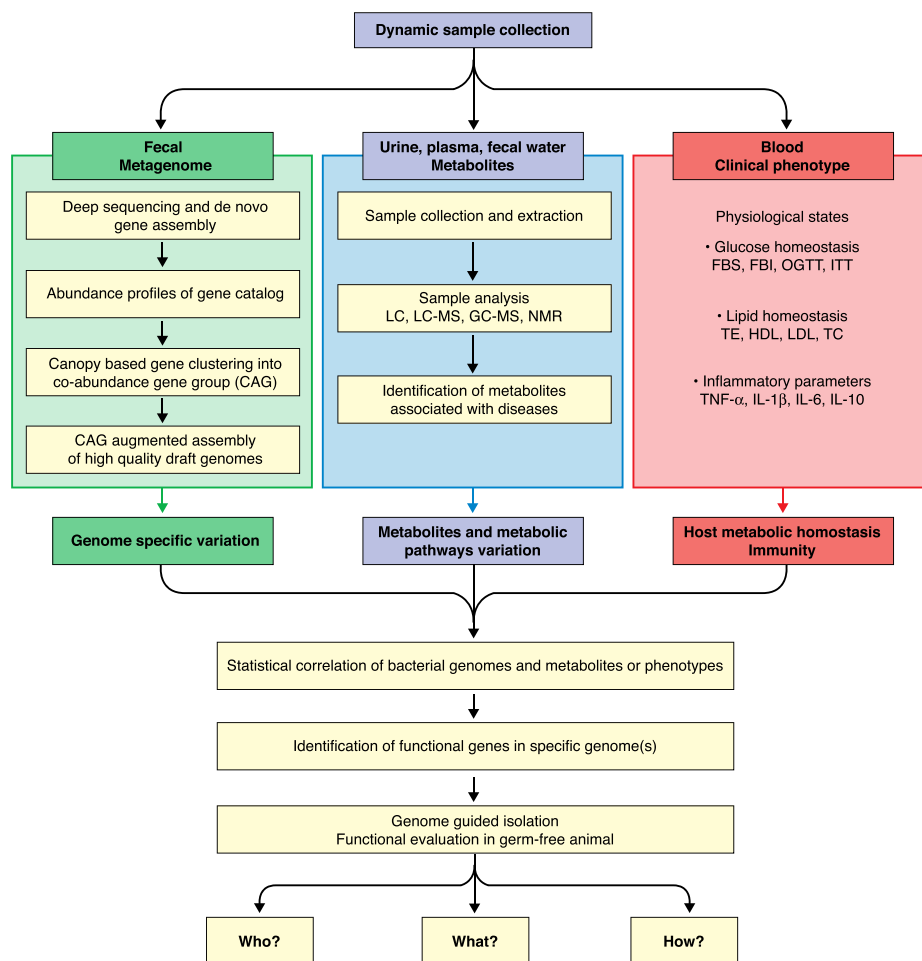


表 团队成员

赵立平	生命科学技术学院教授、美国微生物科学院 Fellow
张晓君	生命科学技术学院研究员
庞小燕	生命科学技术学院副教授
张梦晖	生命科学技术学院副研究员
王凌华	生命科学技术学院高级工程师
赵宇峰	生命科学技术学院高级实验师
申 剑	系统生物医学研究院副研究员
张晨虹	生命科学技术学院助理研究员
吴国军	生命科学技术学院博士研究生

已有大量研究报道肠道菌群与人体代谢性疾病紧密相关，但是在阐述两者的因果关系方面仍存在挑战。本课题对多种代谢性疾病（包括小胖威利综合征、二型糖尿病以



及多囊卵巢综合征) 患者进行的膳食干预, 在改善患者代谢水平及生理指标的同时, 具有调节肠道菌群结构和功能的作用。我们依托  $\pi$  高性能计算及大存储空间的资源, 利用生物信息学和生物统计学的方法, 通过对海量多组学数据 (包括代谢组学和元基因组学等 TB 级原始数据) 分析, 在菌株水平寻找到了部分在疾病发生发展过程中起着重要作用的菌株及菌株之间相互作用形成的功能团, 并发表了相关论文 (Chenhong Zhang et. al, *Ebiomedicine*, 2015; Guojun Wu et al., *Sci Reports*, 2016)。基于上述结果, 可以通过序列引导下的分离方法, 分离得到关键菌株, 进而进行分子水平的操作及动物实验的验证, 以此最终阐明肠道菌群与代谢性疾病的因果关系。利用已搭建的分析平台, 我们近期已开展针对大规模二型糖尿病人群的以肠道菌群为靶向的膳食干预及机理研究。值得注意的是, 随着高通量测序技术的进步与价格下降, 在菌株水平的深度研究将越来越普及, 这将导致样本量的增加以及测序深度的加大, 所以未来对计算性能及数据存储的需求将会大量增长。

## 5 代码优化与研究

除了提供“管机器、开帐号”这些基础服务之外，我们还向国际一流大学的高性能计算中心学习，结合我们的专业知识，为用户提供代码调优的高级服务。并结合这部分工作，开展了研究，培养了学生。

### 5.1 用户代码调优

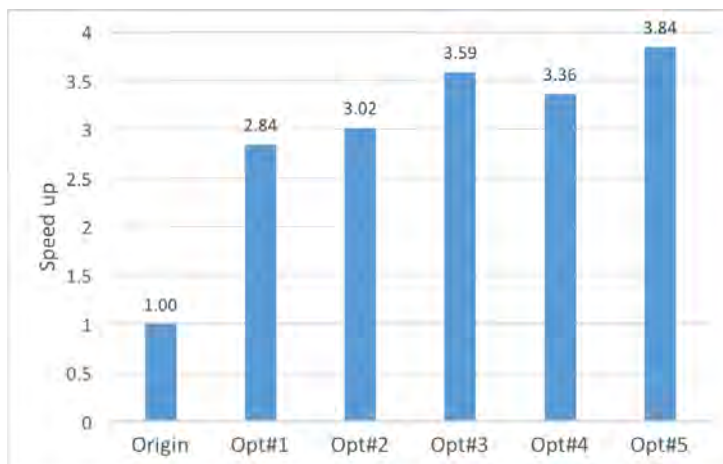
2014 年初，我们推出了 Code Modernization, Acceleration, and Scaling (CMAS) 服务，为  $\pi$  用户免费调优他们的大规模程序。一方面减少程序的运行时间，提高整机利用率，另一方面能让用户在合理的时间内计算更大规模的问题，获得比之前结果更高的精度。CMAS 具体是指：

- **Modernization**：针对 CPU 的并行化和向量化的代码优化
- **Acceleration**：针对 GPU 的移植与优化
- **Scaling**：针对跨节点的大规模并行化

CMAS 服务最主要的挑战是用户习惯于在合理的时间内获得正确的结果即可，因此不太愿意花时间去配合我们调优代码，而更希望把精力放在解决专业问题上。因此只有通过多次充分沟通，建立信任并落实切实可行的技术方案来解决。

到了 2014 年底，我们重新审视这项工作时发现，由于我们团队人手有限，这项工作开展进度没有达到预期。因此 2015 年初，我们推出了 CMAS 2.0 服务，不再仅依靠我们团队的力量，而是充分利用外面资源，构建“用户 + 高性能计算中心 + 外部团队”的三方合作模式。我们不仅在专业用户和外部团队之间起到了纽带和桥梁的作用，并在整个代码调优工作中做出了关键性的贡献。

## 案例一：VLPL 粒子模拟方法优化



VLPL 是基于 PIC (Particle In Cell) 的粒子模拟方法，是从微观角度研究等离子体的某一小区域（例如上千个电子德拜长度）和较短的时间范围（上千个电子等离子体振荡周期）发生的物理过程。等离子体粒子模拟方法的基本过程是这样的：(1) 初始化粒子的位置和速度分布，求出电荷密度和电流密度分布；(2) 数值求解 Maxwell 方程组（用 FFT 方法或空间网格差分法）得到电磁场分布；(3) 由插值得到粒子所在位置处的电磁场；(4) 由单粒子运动方程求出粒子新的位置和速度分布。如此循环下去，就可知道电磁场和每个粒子的运动变化情况。

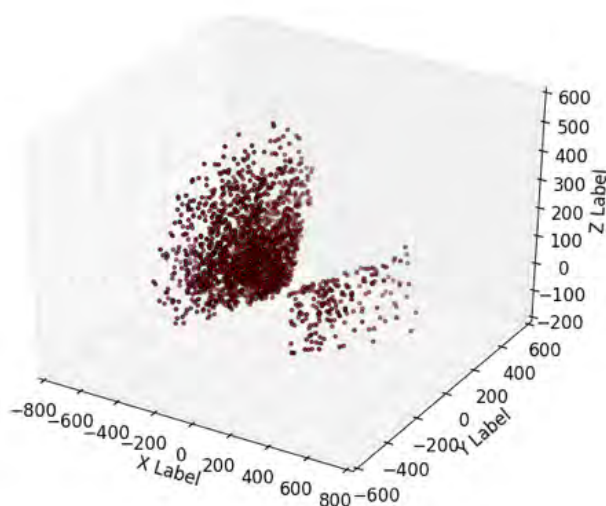
VLPL 程序由交大物理与天文系盛政明课题组陈民老师提供，由高性能计算中心、Intel 公司和物理与天文系合作优化。原程序使用 C++ 编写，并且已通过 MPI 并行化。粒子使用数组储存，每个 Cell 内使用链表储存此 Cell 内的所有粒子。经过分析，原程序的瓶颈在于对场数据和粒子数据遍历时的数据访问，因此优化会主要针对于访存问题。此外，程序的核心部分不存在数据依赖，可以进行向量化以更好的利用 CPU 计算能力。

使用的主要优化方法包括：

- 减少冗余计算：经过对代码和算法的分析，发现存在多余的对整个模拟空间的遍历，将其改为对需要变动的数据的遍历；
- 使用-ipo 编译选项，减少虚函数调用，使用单精度计算；
- 预取：在计算粒子运动时预取下一个粒子，由于粒子使用链表储存，基础的预取只能预取下一个粒子，粒子的预取也会计算的向量化做准备；
- 向量化：计算粒子运动时，将 8 个粒子打包组成 256 位宽的向量，使用向量化指令计算；
- 更远的预取：在 Cell 内保持一个粒子指针的数组，这样在计算时就可以于预取下一个 cell 甚至更远的 cell 的粒子。

经过一系列的优化，在同样测试平台上，VLPL 程序的整体性能获得了 3.84 倍提升。

## 案例二：HBGF 星系聚类算法优化

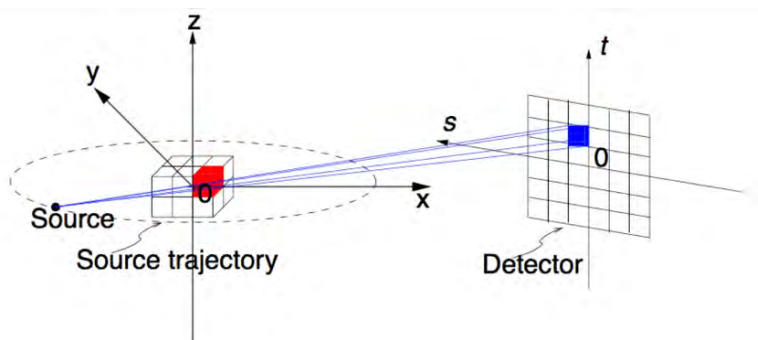


天文学研究中一大重要问题就是对于星系组与星系团的形成与结构的探索。能否合理、有效地分析与利用上述数据，对于此类研究至关重要。Halo-based group finder (HBGF) 算法由上海交通大学杨小虎教授等人提出，是一种基于星系公有暗物质晕的星系成组算法。它可以将每个星系划分到唯一分组，随后人们可以对星系组进行分析，从而研究其空间位置分布以及包括光度、质量、形状以及星系成员等在内的属性。该算法已被用于 2 度视场星系红移巡天，斯隆数字巡天以及各种模拟数据的星系分组，并取得了较好的实验结果。

高性能计算中心与交大物理与天文系的杨小虎教授团队合作，对该应用展开优化。之前在单节点上利用 OpenMP 并行化方法、数据与排序策略等优化手段得到了良好的优化效果，最终使用 12 线程获得程序整体 8 倍的加速效果。

对于更大规模数据集的处理要求更强大的处理能力与内存资源，单计算结点已经无法满足这一需求。采购具有更大内存、更强计算能力，同时价格也更为昂贵的共享内存计算机是一种解决方案，但是若数据集继续扩大并超过其内存容量，则将再次陷入僵局。因此，利用价格相对较为低廉的多 Linux 节点进行并行处理，不仅可以提升程序运行效率，处理更大规模的数据集，而且具有良好的可扩展性。目前的工作方向为使用 MPI 和 UPC 将程序扩展到多节点，并对这两种不同的方法进行比较和性能建模。

### 案例三：SF 医学成像算法在 GPU 上的优化



利用 X 射线计算机断层成像 (X-ray CT: X-ray computed tomography) 技术可以产生人体内部的三维解剖图像。这些图像可以用于各种医学学科中的诊断和治疗。当前大部分商用的 CT 系统使用滤波反投影 (FBP: filtered back projection) 技术来进行图像重建。FBP 技术需要高剂量的 X 射线来产生高质量的诊断图像。而研究表明 CT 扫描使用的 X 射线所产生的电离辐射足以直接或间接地损害 DNA, 将患者置于辐射诱发癌症的危险之下。随着 CT 系统在医学领域的迅速普及及广泛应用, 降低辐射已经逐渐成为公众和医学界关注的焦点。基于模型的图像重建 (MBIR: model-based image reconstruction) 技术可以很大程度地降低 X 射线的辐射剂量, 并且提供高质量的图像。然而, MBIR 技术有一个很大的缺点, 它需要 30 到 60 分钟才能处理好扫描的数据以及重建出一张满足诊断质量要求的图像。而传统的 FBP 技术, 这个过程只需要几秒钟。

MBIR 技术的主要瓶颈在于前向投影 (Forward projection) 和反向投影 (Back projection)。分离足迹法 (SF: separable footprint) 是正反向投影的一种计算方法, 它是比以前工业标准距离驱动 (DD: distance-driven) 方法更先进的技术, 由于其内在的可分离性质, SF 方法可以充分利用并行计算技术进行加速, 从而大大提高 MBIR 技术的速度。本项目研究的是如何利用图形处理器 (GPU: graphics processing unit) 来对 SF 方法进行加速。

本项目跟美国密西根大学 Jeffrey Fessler 教授课题组合作, 交大高性能计算中心与交大密西根学院龙泳老师共同对 SF 方法进行 CUDA 代码优化, 在旧版 CUDA 代码的基础上, 实现了 SF 算法的新版 CUDA 代码。相比于旧版的 CUDA 代码, 同样平台下新版 CUDA 代码对正反向投影的加速约为 2.2 倍。而相比于串行 CPU 代码, 加速比可以达到 24 倍左右, 加速效果相当明显。

## 5.2 编译选项优化

除了代码级的优化，高性能计算中心还在编译选项、并行库、数学库上做组合优化，为用户提供了性能优异的库和应用程序。与用户自行编译的版本相比，经过优化编译的版本无论在稳定性和稳定性上都有很大提高。以交大物理与天文系杨小虎教授 L-Gadget 仿真程序为例，2014 年在  $\pi$  超算上首次使用 2048 核完成了  $3072^3$  粒子系统的仿真，耗时 14 天。2015 年经过编译选项调优，同一类型的仿真耗时缩短到 6 天，且运行过程中没有发生异常中断。目前，经过编译选项优化的库和软件包括 FFTW、MVASP、OpenBLAS、Gromacs 等等，涵盖超算用户的主要学科。在用户有需求、软件授权许可的情况下，高性能计算中心将继续推进这项工作。

## 5.3 代码优化研究

为了给 CMAS 服务提供可持续发展的动力，并为交大培养相关领域的人才，我们针对关键用户的关键应用，围绕  $\pi$  集群，在计算机系统结构及代码优化 (Computer Architecture and Code Optimization) 上开展了一系列的研究。由于  $\pi$  集群的一半的计算能力来自于 NVIDIA GPU 加速卡，为了使用户代码能够在 GPU 以及下一代 Intel Xeon Phi 上获得最佳的性能，我们对它们的体系结构进行了深入的研究，熟练掌握了相关的代码优化手段。基于这些研究工作，我们发表了学术论文，并新申请到了美国 NSF 的课题。

## 学术论文

1. 林新华, 秦强, 李硕, 文敏华, 松岗聪. 使用 Stencil 评估 Intel AVX2 vgather 指令. HPC China 2015. 中国无锡
2. 韦建文, 许志耿, 王丙强, Simon See, 林新华. 异构集群上的宏基因组基因聚类优化. HPC China 2015. 中国无锡
3. 郝赫, 司雨濛, 韦建文, 文敏华, 林新华. 天体物理成团研究中的非规则访存优化. HPC China 2015. 中国无锡
4. 许志耿, 邱君仪, 韦跃明, 董一哲, 韦建文, 林新华. Intel Xeon Phi 平台上 HGGF 程序实现与优化. HPC China 2015. 中国无锡
5. Suttinee Sawadsitang, James Lin, Simon See, Francois Bodin, Satoshi Matsuoka. Understanding Performance Portability of OpenACC for Supercomputers. International Parallel and Distributed Processing Symposium Workshop (IPDPSW) 2015. Hyderabad India

6. Wengqiang Li, Guanghao Jin, Xuwen Cui, Simon See. An Evaluation of Unified Memory Technology on NVIDIA GPUs. Cluster, Cloud and Grid Computing (CCGrid) 2015. Shenzhen China
7. 林新华, 李硕, 赵嘉明, 松岗聪. Intel Knights Corner 的结节点级内存访问优化. 计算机科学 2015
8. 王一超, 秦强, 施忠伟, 林新华. 在 Intel Knights Corner 和 NVIDIA Kepler 架构上 OpenACC 的性能可移植性分析. 计算机科学 2015

## 科研项目

1. 顾一众、林新华等, 863 项目, 高性能计算环境应用服务优化关键技术研究, 2014AA01A302, 2014-2016
2. 林新华、王一超等, 美国 NSF SAVI 项目, A Model for Performance Evaluation of HPC Codes on GPU systems, 2015-2016



## 6 人才培养与培训

高性能计算人才可分为两种，“做超算”和“用超算”的。中心通过与计算机系联合培养高性能计算方向的学生，以竞赛培养这些学生对高性能计算的兴趣；同时也通过暑期课程、系列培训和专项奖学金等多种方式鼓励更多的学生来用超算。

### 6.1 研究生培养

在校领导和研究生院的支持下，中心与计算机系合作，联合培养高性能计算方向的学生。2015 年委托计算机系招收了 4 名研究生，2016 年将毕业 2 名研究生。

### 6.2 竞赛

为了鼓励计算机系学生积极投身高性能计算领域，激发他们的学习兴趣，中心每年组织学生参加相关竞赛。由中心组织的队伍参加了 2015 年 ASC 世界大学生超算竞赛，顺利进入决赛并获得应用创新奖；2015 年 9 月份，中心组织两支队伍依托交大校内应用，分别进入全国并行编程挑战赛最佳应用组和并行优化组决赛。在 11 月无锡举办的决赛表现良好，双双获得铜奖。

表 历年学生竞赛成绩

年份	获得奖项
2012	第 3 届全国 CUDA 编程大赛三等奖
2013	亚洲大学生超算比赛 (ASC13)，首次进入决赛 第 1 届全国 RDMA 编程大赛三等奖
2014	世界大学生超算竞赛 (ASC14) 总冠军 第 2 届 Intel 全国并行编程挑战赛并行优化组铜奖 第 2 届全国 RDMA 编程大赛三等奖
2015	世界大学生超算竞赛 (ASC15) 应用创新 第 3 届 Intel 全国并行编程挑战赛最佳应用组铜奖 第 3 届 Intel 全国并行编程挑战赛并行优化组铜奖



## 6.3 课程与培训



为了提高交大师生高性能计算的应用水平，普及高性能计算知识，高性能计算中心从 2010 年开始，坚持每年举办 6 场以上的 HPC Seminar 系列讲座，2015 年以来较之以往丰富了讲座形式，主要分为以下 3 种：

- 邀请海内外著名学者来交大介绍本领域的前沿发展，如 3 月份邀请了美国 Argonne 国家实验室计算机科学家 Pavan Balaji 博士开展 MPI 编程培训、5 月份邀请到上海纽约大学张峥教授为大家分享深度学习领域的前沿技术；
- 邀请工业界的知名专家介绍高性能计算领域的前沿技术，中心员工参与讲解实例，比如 6 月份邀请美国 PGI 公司王珍博士主讲 OpenACC 并行程序开发，中心员工负责讲解 CUDA 编程基础；
- 由中心员工主讲，针对  $\pi$  用户的基础培训，介绍集群的操作方法以及常用软件环境的配置等。比如 1 月份在密西根学院举行的用户培训和 7 月份在中心举办的  $\pi$  用户现场操作培训。

表 2015 年 HPC Seminar 系列讲座

月份	主题
1 月	Getting Performance Boost Without Code Modification
3 月	MPI 编程培训
5 月	GPU 加速深度学习研讨会
6 月	OpenACC/CUDA 编程及 PGI 编译器讲座
7 月	$\pi$ 集群用户基础培训
10 月	Intel Code Modernization 知识讲座

## 6.4 专项奖学金

为了鼓励交大的博士生和硕士生在  $\pi$  超级计算机上进行高性能计算相关的研究,在校基金会的支持下,中心于 2012 年与 AMD 公司联合设立了“上海交通大学-AMD 高性能计算奖学金”,由 AMD 公司每年出资 5 万元对 5 名获奖者进行奖励。由 3 名校内专家、1 名上海超算中心专家及 1 名 AMD 专家担当评委。2012 年首次举行有 4 人获奖,2013 年 6 人,2014 年 5 人,2015 年 5 人。历年获奖的学生名单如下:

表 历年专项奖学金获得者名单

年份	获奖者	院系	导师
2012 年	刘晋陆	物理与天文系	盛政明
	周 洋	船建学院	张景新
	徐 锋	电信学院	李明禄
	卞亚涛	电信学院	刘允才
2013 年	黎飞宇	物理与天文系	盛政明
	李 平	物理与天文系	罗卫东
	王昱焜	生命科学与技术学院	魏冬青、胡丹
	贾 犇	生物医学工程学院	韦朝春
	王一超	电信学院	李明禄
	敬 超	电信学院	李明禄
2014 年	朱文欢	物理与天文系	董 兵
	曾 明	物理与天文系	盛政明
	干则成	数学系	徐振礼
	邵 成	密西根联合学院	鲍 华
	秦 强	电信学院	李明禄
2015 年	胡智强	生命科学与技术学院	韦朝春
	梁逸浩	物理与天文系	邢向军
	孟德龙	电信学院	林新华、李明禄
	谭海娜	化学化工学院	周永丰
	许志耿	电信学院	林新华、李明禄





结合学校国际化的战略和中心自身发展的需求，我们积极推动与世界一流大学高性能计算中心开展实质性合作，为交大计算相关的科研实力提升提供坚实保障。2015年，我们与 Princeton 大学联合申请了美国 NSF 的 SAVI 项目，与日本高校的主要超算中心（如东京工业大学的 GSIC）保持了紧密联系，与澳大利亚国立大学在教学与学生培养上开展了交流。此外，还连续第 5 年获得了 NVIDIA GPU 卓越中心的支持。

## 7.1 Princeton 大学计算科学与工程研究院

Princeton 计算科学与工程研究院 (Princeton Institute for Computational Science and Engineering, 以下简称 PICSciE) 成立于 2002 年，是国际上著名的以计算为中心的跨学科合作的研究所。2015 年，我们与 Princeton 的国际化合作有了实质性的进展。

PICSciE 的执行理事 William Tang 教授曾多次访问交大，希望交大能借鉴 Princeton 建立 PICSciE 中的经验与教训，逐步推动校内以计算为基础的跨学科合作。2015 年初，Tang 教授受聘为中心的杰出访问教授 (Distinguished Visiting Professor)。5 月，Tang 教授再次访问交大，商定了申请书中的关键细节。此后与我们联合向美国 NSF 提交了 SAVI (Science Across Virtual Institutes) 项目申请，并于 11 月获得了 NSF 的批准。这是国内高校第一个成功申请到的旨在提升跨学科合作能力的 SAVI 项目。

该项目的目标是开发大规模并行应用 GTC-P 的 OpenACC 版本，研究 OpenACC 编程语言用于真实应用时的性能可移植性。GTC-P (全称 The Princeton Gyrokinetic Toroidal Code) 是由 Princeton 大学激光物理实验室 (Princeton Plasma Physics Laboratory, 简称 PPPL) 开发的应用于模拟托卡马克装置内激光湍流的大规模并行代码。该应用的基本方法是利用 PIC (Particle-In-Cell) 算法求解 Vlasov-Poisson 方程，从而对于离子通过托卡马克装置时的运动状态进行模拟。GTC-P 不仅具备重要和实际的物理意义，而且由于其优良的可并行性，被美国能源部下属的 NERSC 超算中心选为官方

应用测试程序，用于对超级计算机的整机性能进行评估。

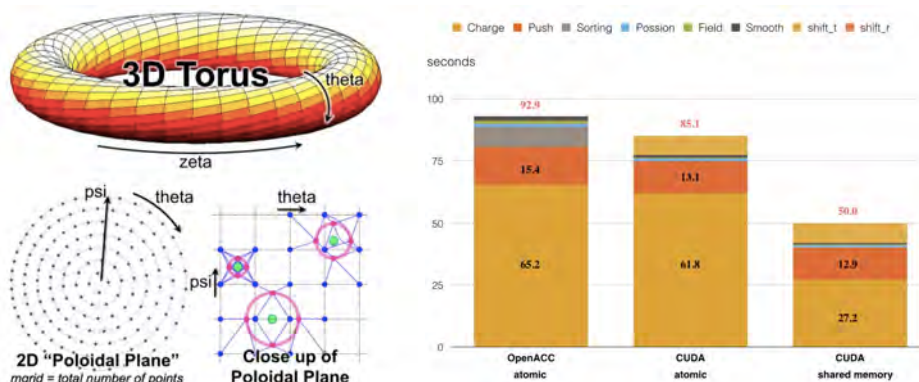


图 托克马卡装置 (左) 和 GPC-P 代码不同版本的性能 (右)。

OpenACC 是一套基于制导语句形式的并行编程语言标准。通过为用户提供脱离硬件设计相关的功能接口，OpenACC 满足了用户编写一套代码就能在多个硬件平台上实现并行化的需求。OpenACC 版本的实现将使 GTC-P 代码能在不同设备间保持可移植性，避免对不同硬件环境需要开发不同版本的问题。

目前我们已经在单节点上完成了 OpenACC 版本 GTC-P 的开发和优化，并在 NVIDIA K80 GPU 上对于该版本进行了性能测试及评估，结果如图。由于 OpenACC 尚未能有效支持 GPU 上 shared memory 的使用，因此我们利用了原子操作来保证多进程间通信的正确性。该版本比同样使用原子操作的 CUDA 版本仅慢了 8%，但比充分利用了 shared memory 优化的 CUDA 版本却慢了 40% 多。下一步，我们计划将 OpenACC 版本的 GTC-P 代码扩展到多 GPU 卡上。

## 7.2 日本高校的超算中心

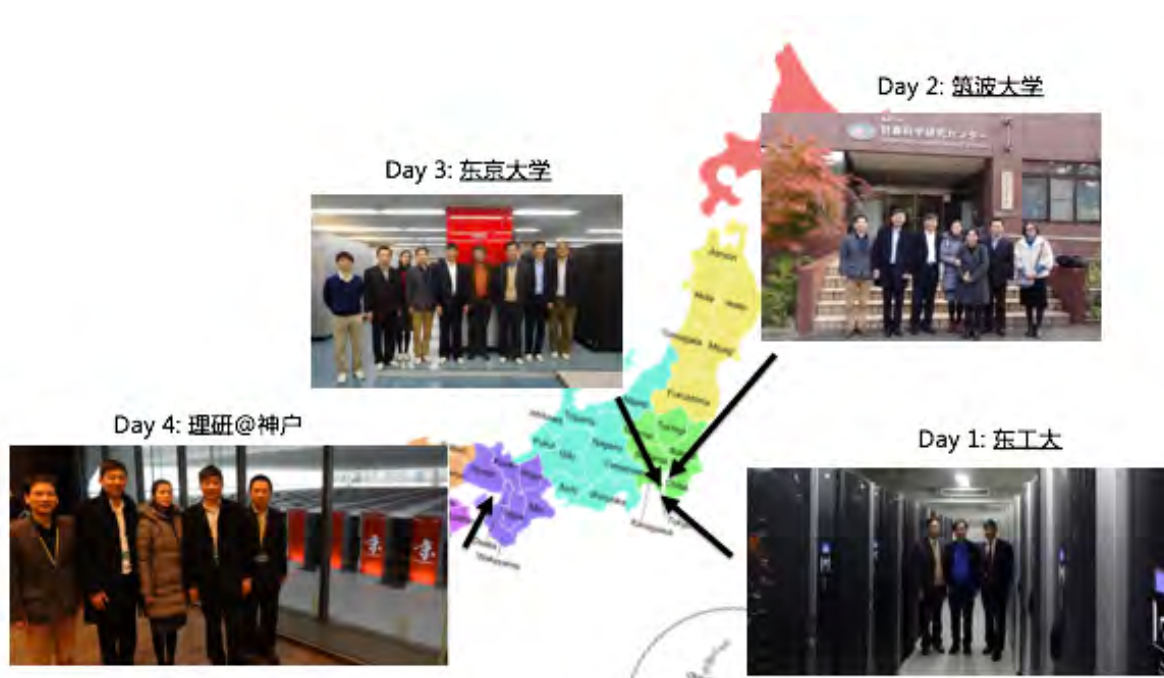
为了更好的规划中心作为校级超算中心的发展路线，学习国际一流大学在这方面成熟的经验，加强与亚洲各高校在高性能计算领域的联系和合作，12 月中心派团访问日本。

第一站是日本东京工业大学，它拥有全日本高校第一的超级计算机 TSUBAME。国际学术情报与计算中心 GSIC (Global Scientific Information and Computing Center) 主任 Isao Yamada 教授和国际著名高性能计算专家、ACM Fellow Satoshi Matsuoka 教授接待了访问团，并介绍了 GSIC 的发展历史以及目前所承担的运维任务和科研工作。访问团一行参观了 TSUBAME2.5 超级计算机和能耗比世界 (Green500) 第一的 TSUBAME-KFC 浸没式油冷原型机，并对交大网络信息中心与东工大 GSIC 两个中心

之间签订合作备忘录 (MoU) 进行了积极探讨。

第二站和第三站分别访问了筑波大学和东京大学，这两个大学计划联合搭建新一代的超级计算机 Post-T2K。筑波大学的 Taisuke Boku 教授和东京大学的 Kengo Nakajima 教授分别介绍了各自学校在高性能计算领域的发展历史以及 Post-T2K 未来的规划。

最后一站是位于神户的理化研究所的 AICS，这里部署了全日本第一的超级计算机“京”。高性能计算编程框架研究团队的负责人 Naoya Maruyama 博士接待了访问团，介绍了 AICS 的概况及所领导团队从事的研究工作，并陪同参观了“京”。此外双方还对一些共同感兴趣的研究问题提出了合作意向。



### 7.3 澳大利亚国立大学计算机研究院

澳大利亚国立大学计算机研究院 (Research School of Computer Science) 的主任 Alistair Rendell 教授于 6 月访问交大，针对高性能计算的人才培养和教学与我们进行了充分交流。之后该学院的 Peter Strazdins 教授于 11 月在交大进行了为期三周的学术访问。Strazdins 教授曾担任澳大利亚国立大学计算机研究院分管教学的副主任，在高性能计算教学上有着丰富经验和较高的国际影响力。中心协助 Strazdins 教授成功申请了 2016 年交大暑期小学期课程，为校内对高性能计算感兴趣的师生提供了一次很好的学习机会。



图 Rendell 教授(左)、Stranzdins 教授(右)访问交大。

## 7.4 NVIDIA GPU 卓越中心

GPU 卓越中心是国际上对一所高校在 GPU 计算领域综合实力的最高认可。目前国际上仅有 24 个，且绝大多数是世界一流名校。英国只有剑桥大学和牛津大学入选。美国的哈佛大学和斯坦福大学亦在内。

上海交大于 2011 年底成为全球第 16 个 CUDA 卓越中心。借助 GPU 卓越中心这个平台，我们与一些世界一流大学（如剑桥大学）的高性能计算中心开展了多项交流。



图 NVIDIA 创始人、CEO 黄仁勋博士(左)与中心员工文敏华

## 8 用户反馈与建议

### 8.1 用户反馈

用户	院系	用户反馈
陈民	物理与天文系	目前 $\pi$ 上的计算资源、存储资源和软件资源基本能满足我们的需求。高性能计算中心提供的用户支持非常有效，反应迅速。提交资源申请报告的方式也很方便，用户委员会评审的过程透明。
赵一雷	生命科学技术学院	Maintenance of PI is a bit frequent; The wechat channel is good. Reply from the support team is also fast.
杨小虎	物理与天文系	经过 $\pi$ 团队的大力支持，尤其是韦建文的不懈努力，我们顺利完成了 $3072 * 3$ 粒子的数值模拟。为我们后续研究、运行同等规模的数值模拟积累了经验。同时，我们也在开始测试在 $\pi$ 上运行流体数值模拟的性能。在此感谢 $\pi$ 团队的大力支持！
胡丹	自然科学研究院	HPC 团队为我们的科研带来了极大的方便，非常感谢！
金朝晖	材料科学与工程	$\pi$ 团队已具备国际领先的专业水准，团队协作能力不断提高，骨干人员都非常敬业，保证了集群正常运转，为用户提供了高品质服务，难能可贵。



用户	院系	用户反馈
何 峰	物理与天文系	计算和存储资源满足需求，作业的等待时间也在接受范围内，计算过程中偶尔遇到集群故障， $\pi$ 的团队都能快速解决。
罗正鸿	化学化工学院	超算团队很尽责、辛苦，为我们提供了较好的用户保障。
Jakob	自然科学研究院	Pi will become important again upon future upgrades, incl. Xeon E5 with AVX2 (>Haswell), and when the gromacs slow-down problem is fixed.
徐见容	医学院	存储等能满足需求，作业等待时间适当，用户支持非常到位，反应迅速。
陈海峰	生命科学技术学院	HPC 提供的用户支持比较有效，反应比较迅速，用户委员会的评审过程比较透明。
孙 淮	化学化工学院	计算资源、存储资源和软件资源基本满足需求，用户支持反应快并且有效。
徐振礼	自然科学研究院	1. 目前 $\pi$ 上的计算、存储和软件资源已足够我们课题组的需求。2. 2016 年初虽然由于气候问题， $\pi$ 一段时间无法使用，但中心解决问题还是比较迅速有效的，也一直通过邮件和微信群等一系列渠道让我们 $\pi$ 用户知道超算中心维护的进展。3. 提交资源申请报告的方式越来越方便，用户委员会的评审也非常透明，公平、公正。
韦朝春	生命科学技术学院	高性能计算中心的支持非常有效，反应非常迅速，服务质量高。我们非常希望高性能计算中心能够尽快升级，继续提供高质量的计算服务。
赵立平	生命科学技术学院	计算和存储资源一般都能满足要求，尤其是在和超算的团体交流和提出相关要求之后，能够得到积极的回应和支持。

## 8.2 用户建议

用户	院系	用户建议
陈 民	物理与天文系	1. 希望能解决作业等待的时间较长的问题。 2. 希望机器的稳定性还能够更稳定一些, 尽量减少中间的维修时间。
赵一雷	生命科学技术学院	1. Some nodes seem not in good state, because same job runs notably slower on these nodes than others. 2. Often there are small problems causing jobs cancelled.
程 真	环境科学与工程学院	1. 作业等待时间有时过长, 例如本课题作业一般需要占用 16 核时, 高峰期等待时间过长。2. hpc 机器或空调系统等出现问题时建议有备用方案, 以免影响作业运行。3. 建议多组织一些 hpc 平台应用及相关知识培训或讲座内容, 方便非计算机专业学生快速掌握 hpc 平台使用方法。4. 并行库及编译器若更换版本或更改地址建议发送邮件通知用户, 并注明各版本是否安装完整, 以免用户使用时造成误解, 或由于所链接并行库位置更改导致软件或模型运行报错。
胡 丹	自然科学研究院	近期 PI 的稳定性较早期有所下降, 可能需要综合探索其中的原因。
金朝晖	材料科学与工程	期待 Pi 二期尽早上线。
何 峰	物理与天文系	在修改程序, 比如说有效并行化现有计算程序的过程中, 如果超算中心能提供技术咨询, 会让我们的计算更快速有效地在 $\pi$ 这个平台上施展。

用户	院系	用户建议
张斌	航空航天学院	1. 希望能够在 <code>putty</code> 的登陆界面写出经费-机时使用情况。2. 希望有一个简单的输出是我们能够看到 <code>job</code> 的执行目录。虽然我们知道 <code>bjobs -l</code> 能够看到这一条目录，但是太乱了。尤其是当 <code>job</code> 很多的时候，很难找到想要 <code>job</code> 的目录
周永丰	化学化工学院	目前 $\pi$ 上的计算资源、存储资源和软件资源可以满足我们的需求。但是经常出现的情况是超算的服务器经常因为这样或者那样的问题重启或者停止使用，我想在其他超算中心肯定不存在这样的问题。所以，是否考虑在服务方面增加更加专业的人员进行维护。
罗正鸿	化学化工学院	1. 希望下一步能扩大存储空间。2. 用户委员会奖励的用户数过少。3. 我们课题组目前主要的需求是小核时的持续稳定运行。
Jakob	自然科学研究院	I like Pi, however, the gromacs slowdown problem means I haven't used it for production runs in over a year. Students/post-docs still use it, but all important sims are now performed on other clusters.
张健	医学院	因为我们课题组是跑长时间的分子动力学模拟，需要存储大的轨迹，需要大的存储空间。目前，账户只给了 200G 的空间，明显感到存储空间不足。希望能够增加到 500G 以上。
徐见容	医学院	提交资源申请方式可以更加优化，资源使用量的计算可以更加人性化和透明。
陈海峰	生命科学技术学院	现在作业排队等待时间非常长，而且运算速度具有明显的下降，现有的计算资源、存储资源不能满足研究需求。

用户	院系	用户建议
罗卫东	物理与天文系	1. 增加长期存储空间, 可以考虑添加大容量硬盘系统。2. 对计算程序出现的问题, 希望能提供更好的技术支持。3. 提供同一个研究组内数个分立的账号, 不同学生分开使用账号。4. 对于常用的一些开源软件和商业软件, 提供编译好的各种版本, 用户不必自己编译。对于已经取得商业软件版权的研究组, 经确认后可以直接使用。可以借鉴 NERSC 的做法。
孙 淮	化学化工学院	偶尔作业等待时间过长, 去年临时停机次数多, 能否提高集群运行的稳定性。
韦朝春	生命科学技术学院	1. 对本课题组而言, 目前 $\pi$ 最大瓶颈是存储资源。另外, 已经有校内用户提出了两年内 PB 级的存储需求。我们迫切希望学校能够加速升级存储资源以满足校内飞速增长的大数据存储及计算分析需求。2. 每个教师开一个账号的做法希望能尽快改进。理想状况是一个教师开一个组账号, 可以给组内每个学生开设一个用户账号。这样, 一旦发现问题, 可以及时明确到个人。目前的情况是, 发现某个账号有问题, 而这个账号往往是一个课题组共用的, 无法即时明确是哪个学生/老师的操作导致的问题。
赵立平	生命科学技术学院	1. 就资源方面当然是越多越好, 存储似乎有些吃紧了。2. 提交资源申请的报告方面, 如果可以建一个网页系统, 把相关信息填进去, 就像填表一样, 然后下方有评委会意见和评委会信息的反馈, 可以帮助评审的透明, 也让申请失败的用户知道哪里有问题。3. 另, 当发生限制用户资源的情况时, 请及时通知。



# 上海交通大学 高性能计算中心年度报告 2015

编写：顾一众 林新华 韦建文 王一超 文敏华

封面设计：栾婧怡

上海交通大学高性能计算中心

地址：上海交通大学闵行校区网络信息中心205

邮编：200240

电话：021-34206060-8204

主页：<http://hpc.sjtu.edu.cn>

邮箱：[hpc@sjtu.edu.cn](mailto:hpc@sjtu.edu.cn)