

**国家天文台**  
**“老虎” GPU 高性能计算集群**  
**年度总结报告**  
**(2012)**

国家天文台信息与计算中心

2012 年 12 月

## 1. 背景介绍

中国科学院在财政部专项资金的支持下，成功打造了计算峰值超过三千万亿次的 GPU 高效能超级计算系统，部署在国家天文台、中科院过程工程研究所、地质与地球物理研究所、电工研究所、高能物理研究所、紫金山天文台、计算机网络信息中心等单位。通过院所共同出资方式承建百万亿次系统，并最终聚合成计算能力近六千万亿次的分布式 GPU 超级计算系统，形成院 GPU 中心及各所 GPU 分中心的两级管理体制。

国家天文台 GPU 高性能计算集群称为“老虎”，配备有 85 个计算节点，通过 Infiniband 高速网络连接；每个节点安装了 2 个 4 核的 Xeon E5520 中央处理器，24GB 内存，CPU 理论峰值为 1.5 万亿次；每个节点安装了 2 块 Nvidia C1060 图形加速卡，理论单精度浮点计算峰值约为 160 万亿次。

国家天文台 GPU 高性能计算集群于 2009 年底安装完成，2010 年 4 月完成项目中期验收，2010 年 11 月通过国家级项目验收，开始为国内外用户提供计算服务。

老虎集群的建成使得国家天文台在高性能计算支撑服务业务上实现了零的突破。从 2009 年开始为我台广大科学用户提供高性能计算服务，也已取得了丰硕的成果。在国家天文台特聘外籍教授 Spurzem R. 带领的丝绸之路项目组的推动下，我台 GPU 集群在国际高性能计算领域取得广泛的影响，深化了我台与国际天文与数值计算领域的合作。借助 2012 年 Spurzem R. 入选“外专千人计划”的良好契机，丝绸之路项目组和信息与计算中心计划在 2013 年对老虎 GPU 集群进行升级改造，将 GPU 卡 C1060 换成最新一代的 K20，使集群的计算能力达到千万亿次。系统升级后将更好地支持大规模天文数值模拟计算、海量数据处理等研究课题的开展。

## 2. 总体运行情况

2012 年以来，GPU 集群总体运行稳定。截止到 2012 年 12 月初，实际完成作业数 10412 个，实际作业计算总时间 2878689.25 小时。科学用户涉及台内数十个课题组及台外高校及研究机构。

2012 年，GPU 集群管理团队通过采取以下措施，优化了集群的作业调度，使得集群所有用户都能够公平的获取到作业运行时间从而完成自己的计算任务，

简化了集群的使用，发展了许多的新用户。

1). 完善 LSF 集群作业管理系统，充分发挥 LSF 的作业管理功能，建立多种作业队列，保证用户提交的作业能够公平的调度。

2). 编制了 GPU 集群操作手册，使得新用户能够根据手册很快熟悉环境，提交与运行自己的任务。

3). 自主研发了 CPU/GPU 异构环境的 GPU 资源调度系统，使得在我台 GPU 超级计算系统环境下，CPU 与 GPU 作业都能够稳定运行，互不干扰，完全避免了资源冲突造成的节点死机问题的出现。

4). 自今年 6 月以来，每月发布集群运行报告，向所有用户通报集群总体运维，用户使用机时，课题总体机时等情况。

### 3. 科学论文产出

自 2010 年以来，借助于 GPU 集群提供的计算能力，部分用户公开发表了自己的研究成果，详见表 1、表 2。

表 1 以国家天文台为第一单位发表的文章列表

第一作者	其它作者	文章标题	单位	文献索引号
Long, R. J	Mao, Shude; Shen, Juntao; Wang, Yougang	Made-to-measure galaxy models - III Modelling with Milky Way observations	1. 国家天文台 2. 曼彻斯特大学 3. 上海天文台	<a href="#">2012MNRAS.tmp..215L</a>
Long, R. J	Mao, Shude;	Made-to-measure galaxy models - II. Elliptical and lenticular galaxies	1. 国家天文台 2. 曼彻斯特大学	<a href="#">2012MNRAS.421.2580L</a>
Wang, Yougang	Zhao, Hongsheng; Mao, Shude; Rich, R. M.	A new model for the Milky Way bar	1. 国家天文台 2. 曼彻斯特大学	<a href="#">2012MNRAS.427.1429W</a>
Ali Taani	Luca Naso, Yingchun Wei, Chengmin Zhang Yongheng Zhao	Modeling the Spatial Distribution of Neutron Stars in the Galaxy	国家天文台	<a href="#">2012Ap&amp;SS.341..601T</a>
Spurzem, R.	Berczik, P.; Zhong, S.; Nitadori, K.; Hamada, T.; Berentzen, I.;	Supermassive Black Hole Binaries in High Performance Massively Parallel Direct N-body	国家天文台	<a href="#">2012ASPC..453..223S</a>

	<i>Veles, A.</i>	Simulations on Large GPU Clusters		
<i>Berczik, P.</i>	<i>Nitadori, K., Zhong S., Spurzem, R., Hamada, T, Wang, X.W., Berentzen, I., Veles, A., Ge, W.,</i>	High performance massively parallel direct N-body simulations on large GPU clusters	国家天文台	2nd International Conference on High Performance Computing 2012, Kyiv, Ukraine
李长华	<i>崔辰州, Berczik,P., 薛随健, Spurzem,R.</i>	一个基于共享内存的 GPU 资源调度器的设计与实现	国家天文台	科研信息化技术与应用, 2012,V3(5): 29-33

表 2 以国家天文台为合作单位发表的文章列表

第一作者	其它作者	文章标题	单位	文献索引号
<i>Li Shuo</i>	<i>Liu, F. K.; Berczik, Peter; Chen, Xian; Spurzem, Rainer</i>	Interaction of Recoiling Supermassive Black Holes with Stars in Galactic Nuclei	1. 北京大学天文系 2. 海德堡大学 3. 国家天文台	<a href="#">2012ApJ...748...65L</a>
<i>Wang, J.</i>	<i>Hammer, F.; Athanasoula, E.; Puech, M.; Yang, Y.; Flores, H.</i>	Loops formed by tidal tails as fossil records of a major merger	1. Univ. Paris-Diderot 2. 国家天文台	<a href="#">2012A&amp;A...538A.121W</a>
<i>Just, Andreas;</i>	<i>Yurin, Denis; Makukov, Maxim; Berczik, Peter; Omarov, Chingis; Spurzem, Rainer; Vilkoviskij, Emmanuil Y.</i>	Enhanced Accretion Rates of Stars on Supermassive Black Holes by Star-Disk Interactions in Galactic Nuclei	1. University of Heidelberg 2. Fesenkov Astrophysical Institute 3. 国家天文台 4. Main Astronomical Observatory, National Academy of Sciences of Ukraine 5. Peking University	<a href="#">2012ApJ...758...51J</a>
<i>Khan, Fazee I Mahmood</i>	<i>Preto, Miguel; Berczik, Peter; Berentzen, Ingo; Just, Andreas; Spurzem, Rainer</i>	Mergers of Unequal-mass Galaxies: Supermassive Black Hole	1. University of Heidelberg 2. Department of Physics, Government College University	<a href="#">2012ApJ...749..147K</a>

		Binary Evolution and Structure of Merger Remnants	(GCU), Pakistan 3.国家天文台	
<i>Fiestas, J.;</i>	<i>Porth, O.;</i> <i>Berczik, P.;</i> <i>Spurzem, R.</i>	Evolution of growing black holes in axisymmetric galaxy cores	1.University of Heidelberg 2.国家天文台	<a href="#">2012MNRAS.419...57F</a>
<i>Fiestas,Jose ;</i>	<i>Spurzem, Rainer</i>	Dynamical Evolution of Rotating Globular Clusters with Embedded Black Holes	1.University of Heidelberg 2.国家天文台	<a href="#">2012scel.boo k..235F</a>
<i>Shang-Fei Liu</i>	<i>James Guillochon,</i> <i>Douglas N. C. Lin, Enrico Ramirez-Ruiz</i>	On the Survivability and Metamorphism of Tidally Disrupted Giant Planets: the Role of Dense Cores	1.PeKing University	<a href="#">2012arXiv1211.1971L</a>
<i>Umbreit, Stefan</i>	<i>Spurzem, Rainer;</i> <i>Henning, Thomas; Klahr, Hubert; Mikkola, Seppo</i>	Disks around Brown Dwarfs in the Ejection Scenario. I. Disk Collisions in Triple Systems	1. Max-Planck-Institut für Astronomie Heidelberg 2.国家天文台 3. Universität Heidelberg	<a href="#">2011ApJ...743.106U</a>
<i>Ernst, A.</i>	<i>Just, A.; Berczik, P.; Olczak, C.</i>	Simulations of the Hyades	1. Universität Heidelberg 2.国家天文台	<a href="#">2011A&amp;A...536A..64E</a>

#### 4. 用户分布情况

目前，GPU 集群真正计算的用户为 38 个。其中 25 个为国家天文台职工或博士、博士后。其它用户来自北大天文系、海德堡大学、台湾大学等合作院校。GPU 集群的用户使用机时信息如表 3 所示。

表 3 2012 年时机使用统计

用户名	姓名	单位	课题组/首席	使用总机时	百分比
jlwang	王建岭	NAOC	赵永恒研究员	455432.1	15.82%
yanbin	杨彦宾	NAOC	武向平院士	434115.83	15.08%
liusf	刘双飞	KIAA		403563.64	14.02%
peter	Peter	NAOC	Rainer	300820.43	10.45%
lishuo	李硕	NAOC, PK	Rainer	298342.57	10.36%

yongli	平永利	NAOC	赵刚研究员	287409.42	9.98%
veles	veles	UHD	Rainer	181004.66	6.29%
wufq	吴锋泉	NAOC	陈学雷研究员	168527.89	5.85%
zhong	钟诗言	NAOC	Rainer	84248.62	2.93%
qiran	夏启然	NAOC	Rainer	56472.75	1.96%
sunl	孙磊	NAOC	詹虎研究员	52098.37	1.81%
qwang	王乔	NAOC	詹虎研究员	40163.21	1.40%
luju	陆路俊	NAOC	陆路俊研究员	33011.81	1.15%
jdeng	邓建榕	NAOC	武向平院士	13038.12	0.45%
julin	朱玲	NAOC	毛淑德研究员	11828.3	0.41%
liulei	刘磊	NAOC	Rainer	11328.62	0.39%
Richard	Long, R. J.	NAOC	毛淑德研究员	8000.36	0.28%
rainer	Rainer	NAOC	Rainer	6011.79	0.21%
fengc09	Chao Feng	KIAA		5817.61	0.20%
gkenn	Gareth Kennedy	NAOC	Rainer	5539.04	0.19%
sitetest	王红帅	NAOC	姚永强研究员	5364.65	0.19%
xuele	陈学雷	NAOC	陈学雷研究员	5294.76	0.18%
ygwang	王有刚	NAOC	陈学雷研究员	3808.82	0.13%
yfjin	金逸飞	NAOC	黄茂海研究员	2536.37	0.09%
jlzhang	张建立	NAOC	薛随建研究员	1271.9	0.04%
jongsuk	Jongsuk Hong	韩国首尔大学		1173.5	0.04%
vlfgpu	张墨	NAOC	黄茂海研究员	1070	0.04%
hsiyu	Hsi-Yu Schive	台湾大学		926.66	0.03%
rahimi	Awat, rahimi	NAOC	毛淑德研究员	162.9	0.01%
bialas	bialas	UHD	Rainer	130.12	0.005%
xuyang	徐阳	三峡大学, NAOC	陈学雷研究员	77.48	0.003%
lianqi	Lianqi, Wang	Caltech		70.78	0.002%
student02				12.48	
yiqing	Yiqing Liu	KIAA		7.49	
khan	khan	UHD	Rainer	4.73	
ray	王瑞	NAOC	颜毅华研究员	1	
mzemp	Marcel Zemp	KIAA		0.47	

表 4 按课题组分类统计机时使用情况

课题组/首席	总机时	百分比
Rainer	943903	32.79%
赵永恒研究员	455432.1	15.80%
武向平院士	447153.95	15.53%

台外合计	411572.63	14.30%
赵刚研究员	287409.42	9.98%
陈学雷研究员	177708.95	6.17%
詹虎研究员	92261.58	3.21%
黄茂海研究员	3606.37	1.24%
陆路俊研究员	33011.81	1.15%
毛淑德研究员	19991.56	0.69%
姚永强研究员	5364.65	0.19%
薛随建研究员	1271.9	0.04%
颜毅华研究员	1	0.00%

## 5. GPU 机群运行费用

GPU 集群的主要硬件于 2009 年底完成采购及安装。后来根据实际需求，增加了部分软硬件设施。截止至 2012 年 11 月底，整个 GPU 集群的建设及维护费累计约为 850 万元，如表 5 所示。

表 5 GPU 集群的主要经费投入

项目	用途	费用(万)	经费来源	时间
机房建设费	搭建南楼主机房	90	国台	2009
GPU 集群(自筹)	GPU 集群系统	250	国台	2009
GPU 集群(拨款)	GPU 集群系统	250	中央财政	2009
新加 85 颗 CPU	满足 CPU 类型作业	25	国台	2009
GPU 集群电费		27	国台	2010
GPU 集群作业管理系统 LSF	用户作业的调度管理	30	国台	2010
GPU 集群存储系统	GPU 集群高效并行存储系统	98.3	国台	2011
GPU 集群扩展节点服务器	新增登录节点	4.987	国台	2011
GPU 集群电费		30	国台	2011
机房维护费		5	国台	2012
GPU 集群电费		30	国台	2012
合计		840.3		

## 6. GPU 集群管理建议

为了更好地发挥我台老虎 GPU 高性能计算集群的作用，更合理的分配资源，丝绸之路项目组和信息与计算中心经协商提出以下几点建议。

1). 成立 GPU 集群用户管理委员会，对用户及其课题进行审核，合理安排作业使用时间。

2). 将数据中心和高性能计算设施纳入台公共科研设施进行管理、开放服务和修缮维护，尽量避免各课题组重复建设自己的小集群，提高资源的利用率。

3). 建立高性能计算技术支持团队，推动以 GPU 为代表的高性能计算在天文研究中的应用。